The mouse protein phosphatase inhibitor-1 gene

Lorna J. McLaren

Doctor of Philosophy

University of Edinburgh

2001

Contents

Abstract

Intracellular proteins are phosphorylated by kinases and dephosphorylated by phosphatases. Protein phosphatase inhibitor-1 (I-1) inhibits protein phosphatase-1 (PP-1), thereby increasing the phosphorylated state of proteins in the cell. The initial aim of the work reported here was to determine whether I-1 is a kidney 'stem' cell marker (Svennilson *et al.*, 1995); if so, it would be the first unique marker for these cells. The project involved characterizing the protein coding region of the mouse I-1 gene, determining the I-1 protein expression pattern in the developing mouse embryo and analysing potential promoter elements of the I-1 gene.

A mouse genomic library was screened using a mouse cDNA clone homologous to the published rat I-1 mRNA and two types of clones with positive sequence homology to the rat mRNA sequence were isolated. These overlapping clones were sequenced, and analysis showed that the predicted mRNA and protein sequences contain 513 nucleotides and 171 amino acids, respectively, with both sharing over 95% homology with the known rat sequences. Further comparison of the predicted mouse protein I-1 sequence to those of rabbit, rat and human showed that there is strong homology across species, and that all share motifs which are important for inhibiting PP-1 (a KIQF sequence at amino acid positions 9–12 and a threonine at position 35). The mouse I-1 gene contains seven protein-coding exons which lie in a 7 kb region of DNA on chromosome 15, band F.

Svennilson *et al.* (1995) detected I-1 expression in peripheral metanephric mesenchyme (nephron progenitor) cells in rat sections using RNA *in situ* hybridization. With this in mind, the mouse protein expression pattern was determined using whole-mount tissue, a commercially available anti-I-1 antibody and fluorescent confocal microscopy. The results showed that I-1 is not a 'stem' cell marker in the developing kidney, but is restricted to the peripheral epithelial layer of cells of the kidney i.e. the mesothelium. The epithelial phenotype was confirmed by

the use of laminin and pan cytokeratin antibodies, while a Pax-2 antibody showed that the I-1-positive cells were not metanephric mesenchyme. Further analysis of tissue from the coelomic cavity (lung, liver, heart, intestine and gonad) showed I-1 specificity to the mesothelium in embryonic and adult tissue. There was internal I-1 expression in the developing sex cords of ~E13 testis, but expression was down-regulated as the testis developed. Inhibitor-1 was also expressed in the surface layer of ectoderm. Human lung tissue was analysed for I-1 expression but, since it had high levels of autofluorescence, it was unsuitable for fluorescent microscopy.

A construct containing ~9 kb of 5' upstream sequence from one of the isolated genomic clones was inserted into the promoterless vector pEGFP-1, which will only produce an enhanced green fluorescent protein (EGFP) if the inserted DNA contains a functional promoter. A second construct was built using ~3 kb of upstream genomic DNA. The two constructs were used to transiently transfect the cell lines HEK293 (human embryonic kidney) and 4/4 RM-4 (rat mesothelial) using FuGENE 6 (Roche). The transfection protocol was optimized for each of the cell lines but neither of the two constructs resulted in EGFP-positive cells. However, the results do indicate the direction of future work.

Declaration

I declare that:

(a) this thesis was composed by myself;

(b) the work presented here is my own, except where stated; and

(c) the work has not been submitted for any other degree or professional qualification.

Lorna J. McLaren

Acknowledgements

I would like to thank my supervisors for their constant help and encouragement over the past few years: Dr Jonathan Bard, for giving me the opportunity to work on the thesis; and Dr John Mason, for his advice and endless patience, particularly when I was building the %*$@!^£ constructs.

I would like to thank the following people (in no particular order): Linda Sharp for her help on the confocal microscope; Shelagh Boyle for her expertise in FISH; Jean Flockhart, Margaret Keighren and John Verth for letting me have their unwanted mice; Dr Ian Simpson for the supply of various plasmids; Dr Sula Corbett for sharing the cost of the library; Dr Donald Salter for the supply of human lung tissue; Dr Jamie Davies for advice; Dr Rory Duncan for cells; Dr Mark Barnett and Dr Alasdair MacKenzie for advice on constructs; the staff members of the (former) Anatomy department, and in particular, those along the research corridor; Sue Veitch for providing the exercise classes (for my physical and mental well-being!); Renske Brune and Amanda Carter for empathizing; Joan Burton for being there; and to my family.

Finally, I would like to dedicate this work to Andrew J. Wilson, for his love and support, particularly in my darkest hours. I couldn't have gone on without you. And to Kismet: may the tuna fly free.

Abbreviations

~	Approximately
A	Adenine
bp	Base pair
BSA	Bovine serum albumin
CO_2	Carbon dioxide
C	Cytosine
cm	centimetre
DNA	Deoxyribonucleic acid
E	Embryonic age (day)
EGFP	Enhanced green fluorescent protein
FISH	Fluorescent *in situ* hybridization
G	Guanine
GFP	Green fluorescent protein
g	Gram(s)
h	Hour(s)
H_2O	Water
I.M.A.G.E.	Integrated molecular analysis of genomes and their expression
I-1	Inhibitor-1
kb	Kilobase pairs
kDa	KiloDaltons
L	Litre(s)
L^{-1}	Per litre
M	Mole/Molarity
MCS	Multiple cloning site
min	Minute
$MgSO_4$	Magnesium sulphate
mM	Millimolar
mL	Millilitre
mL^{-1}	Per millilitre
MM	Metanephric mesenchyme
mm	Millimetre
M_r	Molecular weight
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ng	Nanogram
nt	Nucleotide
λ	Bacteriophage lambda
PBS	Phosphate buffered saline
PBSTX	PBS plus 1% Triton X-100
PCR	Polymerase chain reaction
PFA	Paraformaldehyde

PFU	Plaque forming units
PP	Protein phosphatase
RNA	Ribonucleic acid
s	Seconds
TESPA	3-aminopropyltriethoxysilane
T	Thymine
UB	Ureteric bud
μg	Microgram(s)
μL	Microlitre(s)
μL^{-1}	Per microlitre
V	Volts
w/v	Weight (g) per volume (L)

List of amino acids

Single letter code	Abbreviation	Amino acid
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutaminc acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ileu	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparganine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

List of figures

List of tables

Chapter 1
Introduction

1.1. Introduction

The research reported here was undertaken to characterize the mouse protein phosphatase inhibitor-1 gene, and to determine whether it is a candidate stem cell marker of the developing kidney, as suggested by Svennilson *et al.* (1995). This chapter consists of: an anatomical section describing the kidney and its development; a description of kidney stem cells; a brief account of genes expressed during kidney development; and an overview of protein phosphatase inhibitor-1.

1.2. Mammalian renal formation: an overview

The mouse embryo is a widely used model for mammalian development. The main reasons for this include ease of handling and a short gestation period that allows embryological development to be determined within a narrow time-scale. Mouse development can be compared to other higher vertebrates, including humans, although one main difference is the process axial rotation (also referred to as turning) of the murine embryo, resulting in the inversion of the germ layers (see review by Kaufman & Bard 1999).

Kaufman & Bard (1999) reviewed the development of the post-implantation mouse embryo, analysing each organ system in isolation and with reference to gross embryonic development. The murine embryo has a gestation period of approximately 19 days. At around E7·5 (E = embryonic day; E0·5 is defined as midday on the day of plug detection), the embryo undergoes gastrulation where the epiblast cells migrate into the primitive streak and separate the epiblast from the underlying hypoblast. This process results in the formation of the three germ layers: the endoderm, ectoderm and mesoderm (Fig. 1.1, Kaufman & Bard 1999).

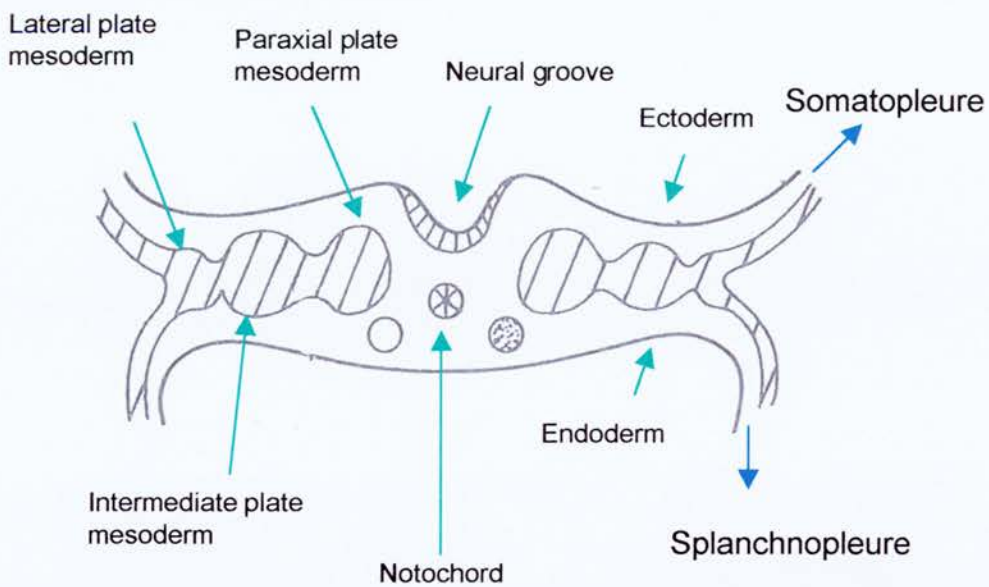


Figure 1.1. Diagram of the developing mouse embryo at ~E8 showing the formation of the mesoderm. The paraxial plate mesoderm forms the somites, the intermediate mesoderm forms the nephrogenic cord, and the lateral plate mesoderm forms the somatopleure and the splanchnopleure (indicated by the blue arrows). Adapted from Kaufman & Bard (1999).

The mesoderm itself partitions into three components:

(1) the paraxial mesoderm which gives rise to the somites at ~E8;

(2) the intermediate plate mesoderm (~E8) which forms the urogenital ridge at E9, and further develops into the gonads, urogenital mesentery, paramesonephric duct and the nephric duct; and

(3) the lateral plate mesoderm which is the last of the mesodermal tissues to form at E8·5, splitting to form the somatopleure (which will form body wall musculature) and the splanchnopleure (which will form gut musculature). The mesothelium is the cellular layer lining the inner coelomic cavities and is derived from this lateral plate mesoderm.

Renal formation in mammals and birds involves inductive interactions between two types of tissue derived from the intermediate mesoderm which form the transient but functionally distinct pronephros and mesonephros before formation of the permanent excretory organ of higher vertebrates, the metanephros. The nephron is the basic functional unit of the metanephric kidney.

The pronephros is derived from the intermediate mesoderm, but both the mesonephros and the metanephros form by condensation of the nephrogenic cord (also referred to as the mesenchymal blastema) that lies in the urogenital ridge. The urogenital ridge itself gives rise to the metanephric mesenchyme (MM) and the epithelial ureteric bud (UB), an outgrowth of the mesonephric or Wolffian duct, and these two tissue types form the metanephric kidney. Although the MM and UB are both derived from the intermediate plate mesoderm, they are otherwise unrelated, neither being descended from the other. In the mouse, the transient pronephros and mesonephros form at E9 and E9·5, respectively, followed by metanephros development at E11 (Kaufman & Bard 1999; Fig. 1.2).

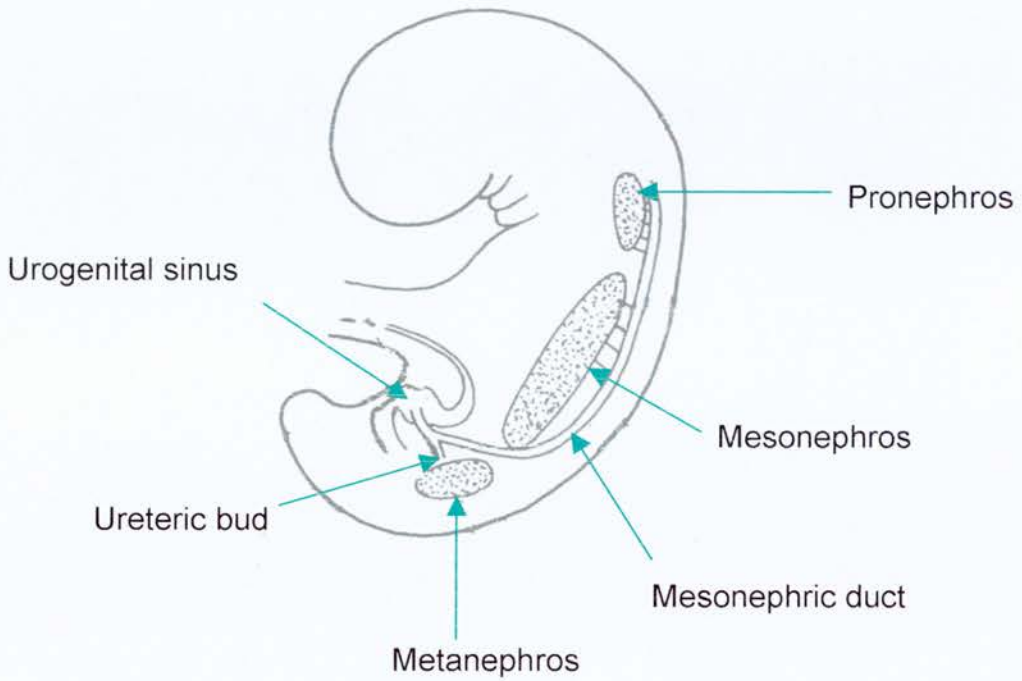


Figure 1.2. Schematic representation of the developing mouse embryo showing the position of the developing pronephros, mesonephros and metanephros. Adapted from Kaufman & Bard (1999).

1.2.1. The pronephros

The pronephros is the functional excretory organ in larval fish and amphibians, but the mesonephros is the permanent organ in the adult organisms. There are differences between the types of nephron found in the pronephros and the more complex excretory organs: the pronephros contains nephrons which are not integrated into the renal tubule (Fig. 1.3). The pronephros of amphibians is a functional structure, but is considered to be one large unintegrated nephron that filters wastes into the coelomic (or body) cavity. In comparison to non-teleost fish and amphibians, teleost fish have a more advanced pronephric structure that consists of a glomerulus, and when two of these structures fuse, a glomus is formed. The pronephros of reptiles, birds and mammals do not form glomera or glomeruli, and may have a limited functional role (for reviews, see Vize 1997; and Kaufman & Bard 1999).

1.2.2. The mesonephros

During the renal development of mammals, the pronephric duct degenerates, but the caudal region forms the Wolffian duct, forming part of the excretory system. The mesonephros is also known as the Wolffian body and is a less complex structure than the metanephros, although both structures have a similar vascular system. Pigs have a functional mesonephros, but the mouse and human structures are obsolete. The mesonephros contains two types of nephron: an integrated glomerulus; and a ciliated nephrostomal connection to the coelom (the latter is often not present in mammals and birds since the placenta is well developed in these organisms; for review, see Kuure *et al.* 2000). As the nephrogenic cord extends caudally, there is regression at the rostral end, and thus, the organ is never fully complete. The paramesonephric (Müllerian) ducts form from the urogenital ridge and eventually become part of the female reproductive system whereas the Wolffian duct forms part of the male reproductive system as the mesonephros degenerates (for reviews, see Vize 1997 and Kaufman & Bard 1999; Fig. 1.4).

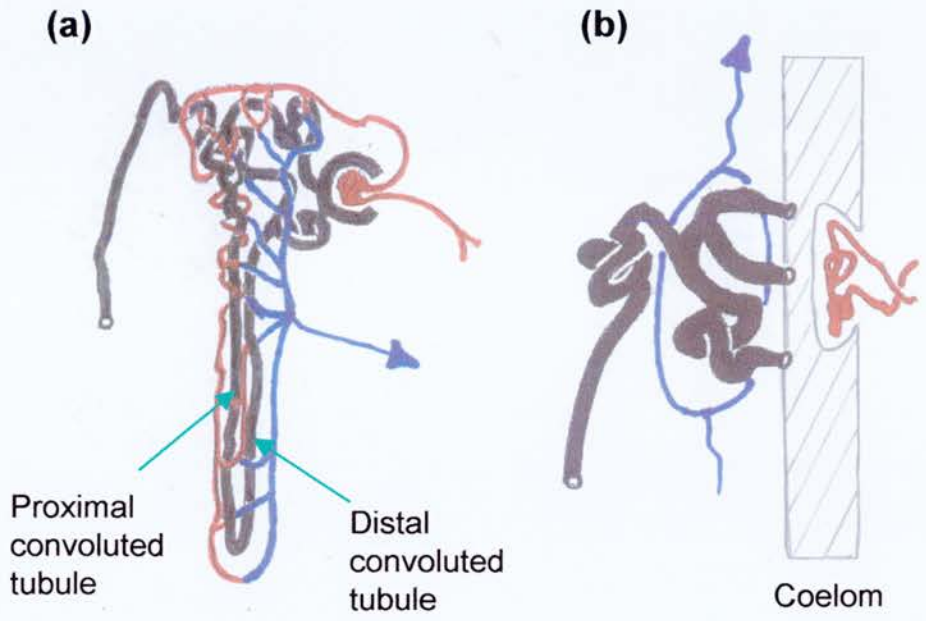


Figure 1.3. The (a) integrated and (b) unintegrated nephron. In the integrated nephron, the blood supply (shown in red and blue to represent entry into and exit from the kidney, respectively) is associated with the proximal and distal tubules of the nephron. In the unintegrated nephron, there is no close contact between the nephron and blood supply. Adapted from Vize *et al.* (1997).

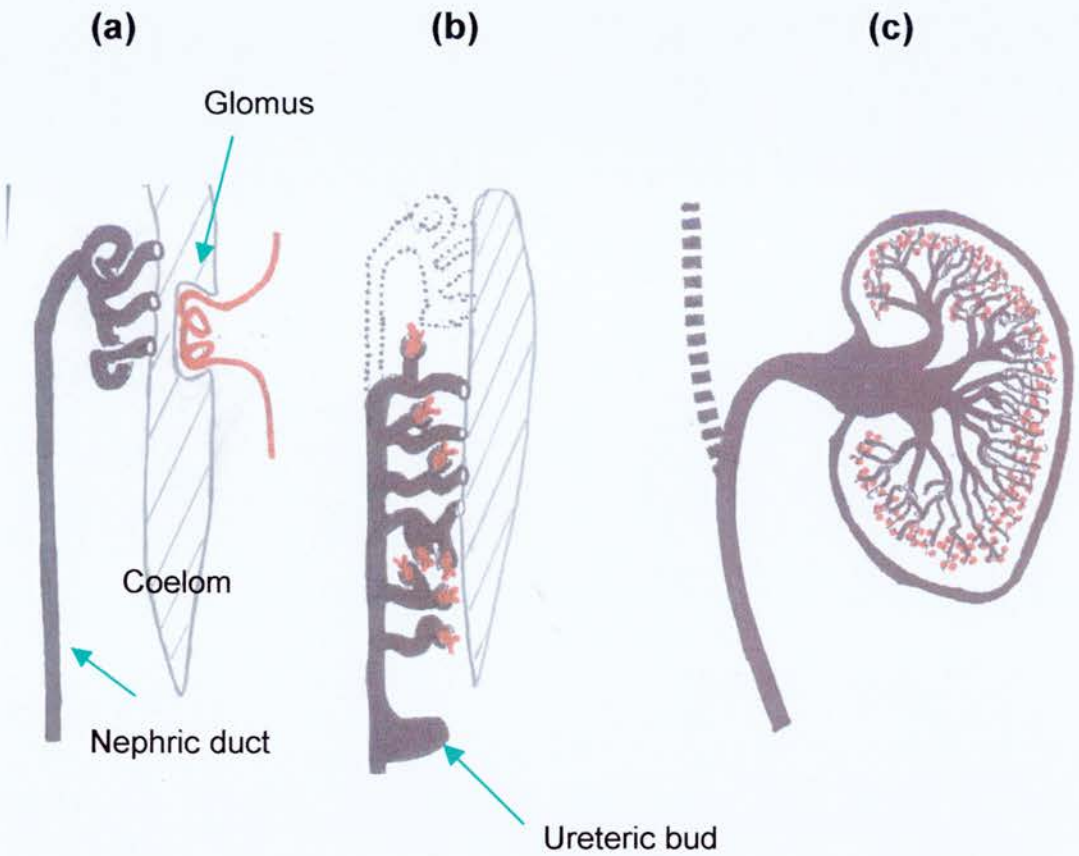


Figure 1.4. Development of the vertebrate excretory system (not drawn to scale). (a) An unintegrated nephron of the pronephros. (b) The pronephros degenerates (dotted lines) as the mesonephros develops, and glomeruli form (10–50 nephrons may be present). (c) The ureteric bud grows from the nephric duct and will eventually form the metanephros, consisting of hundreds of thousands of nephrons. The broken line represents the degenerating mesonephric duct. Adapted from Vize *et al.* (1997).

1.2.3. The metanephros

The metanephric kidney controls osmoregulation of the blood system of higher vertebrates and removes the collected vascular waste; there are approximately one million nephrons in the adult human organ. The maturing metanephros has a gradient of nephron maturation with the fully formed structures situated near the medulla and the early aggregates located at the cortex (Fig. 1.4; for reviews, see Herzlinger 1994; Bard *et al.* 1994; Davies 1996; Vize 1997; Davies & Bard 1998; Kaufman & Bard 1999).

The kidney has been a favoured organ for studying development since its successful culture *in vitro* by Grobstein (1956). This allows the maturation of the organ to be studied and thus, the phenotypic changes which occur *in vivo* can be mimicked *in vitro*. This has obvious benefits to the investigator since it would be difficult to visualize and manipulate such processes *in vivo*. The requirements for normal development of the kidney were also established by Grobstein (1956): the ureteric bud (UB) must be in contact with the metanephric mesenchyme (MM) for elongation and branching, but the MM can be induced by other inducing tissues such as spinal cord. Lack of contact of the MM cells with the inducing UB results in apoptosis of the mesenchyme (Saxen *et al.* 1983), although the survival factors present that prevent apoptosis are unknown (for review, see Davies 1996). However, recent studies have shown that isolated MM and UB can be cultured in the absence of the reciprocal tissue (Karavanova *et al.* 1996; Qiao *et al.* 1999a) although the medium must be conditioned or contain specific molecules (see section 1.5 for details).

Signals from the mesenchyme initiate growth of the UB, which itself induces mesenchyme-to-epithelial transition of the MM to form the nephron (see section 1.5 for details). The altered phenotype associated with mesenchyme-to-epithelial transition is peculiar to the kidney since it is not a feature of other glandular organs, such as the lung and salivary gland (Hay 1995). However, modern techniques, such as gene targeting, have allowed developmental processes to be analysed in terms of genetic controls *in vivo* since genes can be switched off and their role during embryogenesis determined (see section 1.5 for further details).

In the metanephric kidney, the peripheral stem or blast cells of the MM have the potential to form both stromal and epithelial cells. These stem cells are termed pluripotent since they are have one of several fates, as opposed to precursor cells which have a single predetermined fate. During epithelialization, the stem cells condense and form pretubular aggregates, comma- and S-shaped structures, and finally, the nephron. Vascularization of the nephron occurs during the S-shaped phase of development to form the glomerulus and is an important process after the inductive interactions have been initiated (Sorokin & Ekblom 1992). The integrated metanephric nephron consists of the glomerulus and the Bowman's capsule, which are closely associated with each other, as well as the proximal and distal convoluted tubules, and the loop of Henlé. The nephron is derived from the MM and the urinary collecting ducts from the UB (Fig. 1.5; for reviews, see Davies 1996; Vize 1997; Kaufman & Bard 1999).

1.3. Stem cells in the developing kidney

At the end of the nineteenth and in the early twentieth centuries, the renal blast or stem cells were thought to derive from the epithelial UB, suggesting that such cells would have to be included in the MM before nephron development. It is now widely acknowledged that the metanephric blastema is the origin of the range of epithelial cells of the kidney, as highlighted by the work by Grobstein (1956). However, in a review article, Herzlinger (1994) supported the former stem cell theory by postulating that cells of the ureteric bud (UB) can give rise to nephron progenitors within the metanephric blastema. Herzlinger (1994) based this theory on work using a lineage marker system and transfection studies using *lacZ* tagged cells. These studies showed that the uriniferous tubule is derived exclusively from the UB, but also, controversially, that the UB is a source of stem cells for the metanephric blastema. This is at odds with more recent work, particularly with reference to glial cell-line derived neurotrophic factor (GDNF) and its tyrosine kinase receptor, c-ret, which belongs to a family of proto-oncogenes. GDNF is

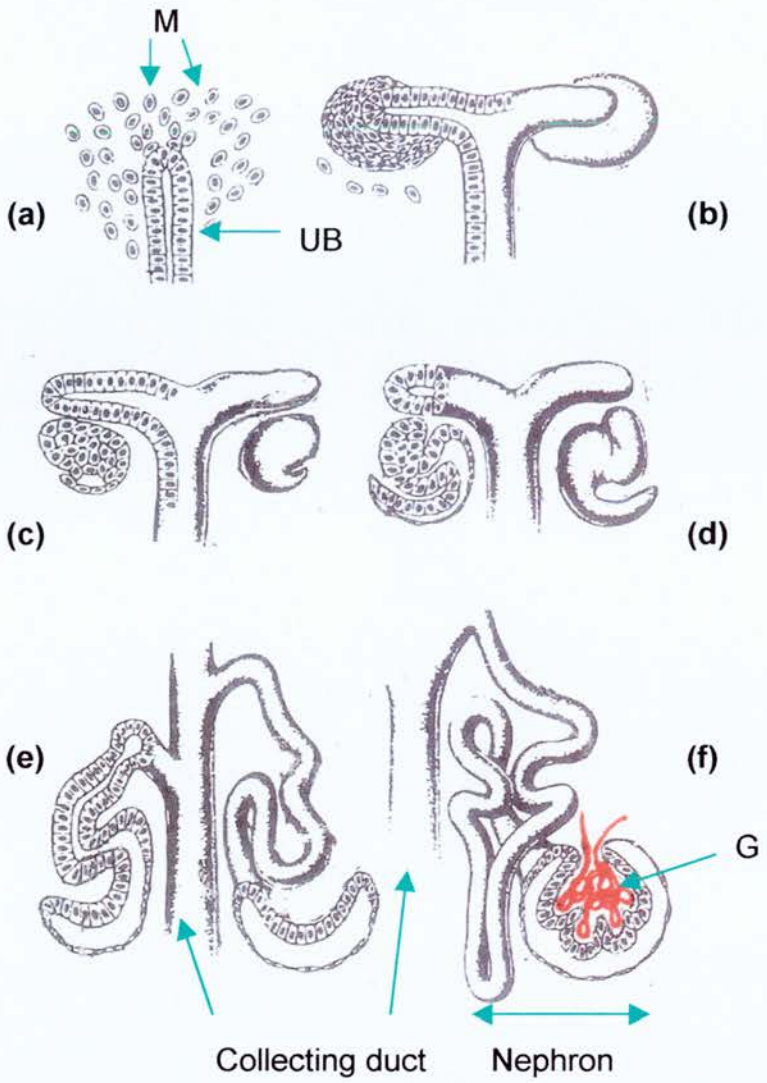


Figure 1.5. Development of the vertebrate integrated nephron: (a) the loose mesenchyme (M) and ureteric bud (UB) undergo reciprocal interactions which lead to (b) condensation of the mesenchyme around the bifurcated ureteric bud; the condensed mesenchyme leads to (c) comma- then (d) S-shaped structures; and (e & f) the final stages of development of the nephron. The nephron is derived from the metanephric mesenchyme and the collecting duct from the UB. Vascularization of the nephron occurs during the S-shaped stage, but the final glomerulus (G) is indicated in (f). Adapted from Sorokin & Ekblom (1992).

expressed by kidney mesenchyme and is required for early induction and subsequent growth of the UB (Pichel *et al.* 1996; Pepicelli *et al.* 1997; Sainio *et al.* 1997). The c-ret receptor is expressed by the UB and disruption of its expression results in abnormal kidney development. The phenotypes of *c-ret* and GDNF knock-out mice are identical (see section 1.5.2 for further details). However, mesenchyme from *c-ret* $^{-/-}$ mice can be induced to form nephrons in culture (Durbec *et al.* 1996; Schuchardt *et al.* 1994, 1996), an observation that would not be expected if Herzlinger's (1994) theory is correct. It may be possible that there is integration of epithelial cells from both the UB and epithelialized MM during nephron formation, and this could explain Herzlinger's (1994) hypothesis.

The term 'stem cell' is contentious since researchers have conflicting views as to its exact definition. However, it is agreed that stem cells are not precursor cells, since the former are multipotent whereas the latter are committed to differentiate to a specific cell type. Bard *et al.* (1996) referred to stem cells as already induced MM, and similarly, Davies & Brandli (1997) suggested that stem cells are early condensates or at least are not uninduced MM, having undergone the initial inductive step. Dressler & Douglass (1992) characterized the transcription factor, Pax-2, which is expressed by the condensed mesenchyme and its epithelial derivatives, but the term 'stem cells' was not used by these authors, although these Pax-2-positive cells could be classified as stem cells since they are early condensates. In this thesis, stem cell will refer to the early-induced mesenchyme.

In the developing kidney, the stem cells are located at the periphery of the organ during development since the cells undergo nephrogenesis after induction by the UB, thus giving the characteristic gradient of nephron formation found in the developing kidney. As described earlier, if the MM is not invaded by the UB, then the uninduced cells die by apoptosis (Koseki *et al.* 1992), but the stem cells must have survival factors to prevent this from occurring. Therefore, the stem cells cannot be uninduced MM since they do not die. The lack of data for stem cell phenotype and exact location during kidney development highlights the need for a suitable stem cell marker. Although Pax-2 (Dressler & Douglass 1992) and

hepatocyte growth factor (HGF; Sonnenberg *et al.* 1993) do label uninduced MM cells, there is no unique marker for the kidney stem cells (Davies & Brandli 1997).

1.4. Other cell types in the kidney

The metanephric mesenchyme, epithelial ureteric bud and the stem cells are not the only cell populations found in the kidney. Proper renal function is only possible if the organ has undergone the correct developmental processes to produce a range of different cell types.

Kidney stromal cells are derived from the mesenchyme and are an important source of growth factors for the developing organ to ensure mesenchyme-to-epithelial transformation, and for UB and collecting duct growth. These key roles were highlighted by analysis of brain factor 2 (BF-2; also termed Foxd1) a forkhead/winged helix transcription factor that is expressed in the mesenchyme-derived stromal cells of the kidney. *BF-2*$^{-/-}$ knock-out mice survived for 24 h after birth; the kidneys were small, fused longitudinally and had a small number of large condensates, although the presence of the stromal cells was unaffected in the knock-out mice. The mutant mice had limited branching of the ducts and there was a failure of epithelial differentiation (Hatini *et al.* 1996).

The glomerular basement membrane is produced by the podocytes and endothelial cells (Ekblom 1981; Sariola *et al.* 1984) and maintains the structure of the glomerulus since it is the framework upon which other cells grow. Laminins are major components of this glomerular basement membrane. The laminin $\alpha 5$ mutant mouse had abnormal basement membrane structure resulting in altered glomerular growth and development: the glomerular epithelial cells were not properly arranged, and some of the mutant mice lacked one or both kidneys (Miner & Li 2000).

Endothelial cells present in the kidney are believed to arise from the blood vessels as the organ develops, as suggested by *in vitro* culture of MM that showed a lack

of both endothelial and mesangial cells (Ekblom 1981). Risau & Ekblom (1986) showed that developing endothelial cells secrete angiogenic factors and suggested that these factors increase blood vessel development in the kidney. Breier *et al.* (1992) analysed the expression pattern of vascular endothelial growth factor (VEGF), an angiogenic mitogen that targets vascular endothelial cells. The protein was observed in epithelial cells of the glomerulus, suggesting a role for VEGF during endothelial growth and development. Kitamoto *et al.* (1997) showed that VEGF is an important molecule for glomerulogenesis since newborn mice injected with an anti-VEGF antibody had fewer nephrons than control animals. Endothelial cells also give rise to neurons but the process of innervation of the kidney has not been fully characterized. Renal nerves do effect kidney development and sympathetic innervation function alters during maturation (for review, see Robillard *et al.* 1993). Afferent renal innervation of rat kidneys is more complete by birth since the nerves are already in their target regions, but the efferent nerves mainly develop postnatally (Liu & Barajas 1993). *GDNF*$^{+/-}$ mice have defective enteric innervation, and lack ureteric bud induction and kidney formation (Moore *et al.* 1996; Pichel *et al.* 1996; Sanchez *et al.* 1996) although whether this neuronal growth factor directly affects kidney innervation has not been determined.

Mesangial cells are mesenchymal, smooth muscle-like cells and maintain the structure of the glomerular capillaries in association with the endothelial cells. Mesangial cell development is dependent upon platelet derived growth factor (PDGF-B; Leveen *et al.* 1994) and its receptor (PDGF-Rβ; Soriano 1994; Lindahl *et al.* 1998). GDNF is expressed in the human kidney; it is a growth factor for the mesangial cells and may have a role in the pathogenesis of glomerulosclerosis (Orth *et al.* 2000).

Dendritic cells are involved in the primary immune reaction, and in the kidney they are located in the glomerular mesangium, cortical cells and the endothelia surrounding the convoluted tubules. Gieseler *et al.* (1997) enriched and characterized the dendritic cells from rat kidneys, and found that there is a shared

phenotype with those from the spleen. Therefore, it is likely that these dendritic cells mount an immunological attack during allograft rejection.

1.5. Genetic control of kidney development

Metanephric kidney development involves the time-specific expression of a range of genes to form the final organ and hundreds of genes are already known to express during kidney development. 'The Kidney Development Database' developed by Davies & Brandli (1997), lists the current genes expressed by the developing kidney and can be accessed on the world-wide web at http://golgi.ana.ed.ac.uk/kidhome.html. However, it is important to isolate the genes which have a direct developmental role from the housekeeping genes that are present for normal cellular function.

Genes contain regions of DNA which code for proteins and regulatory functional elements, the latter of which can be used as a tool for studying kidney development and disease. The technology that allows gene transfer *in vitro* has the potential to be applied *in vivo* (for review, see Fine 1996). For example, vectors which have reporter molecules, such as *lacZ*, can be fused with a stretch of genomic DNA containing regulatory elements. Transgenic mice can then be produced and analysed for expression of the fusion protein. Mutation analysis in mice and molecular techniques, such as differential display, have identified new and previously uncharacterized genes switched on during renal development. Rosenblum & Yager (1997) modified the differential display technique by pooling expressed cDNAs from mouse metanephroi and using this as a multiplex probe to screen a total mouse embryo cDNA library. This allowed identification of previously undetermined genes expressed in the induced mouse metanephros.

The genes expressed in the UB or MM which are responsible for normal kidney development can be broadly divided into the groups of transcription factors, signalling molecules, and the receptors and components of the extra-cellular matrix (ECM).

Figure 1.6 shows the complex relationship between genes, transcription factors and external signalling molecules which regulate gene function leading to developmental changes (reproduced by kind permission of Dr Jonathan Bard).

Disruption of function by gene targeting can confirm the contribution of an individual gene to kidney development and allows the role of the gene in embryonic and adult development to be analysed. However, there are genes which must be expressed during two or more stages of kidney development for normal embryonic development. For example, *Pax-2* is required for outgrowth of the UB during early induction of the kidney, but is also important for the spatial orientation of the ureter (Sanyanusin *et al.* 1995). Therefore, it would be misleading to assume that the *Pax-2* gene only has a role during kidney formation since early death of the embryo would not highlight any later function of the gene. Conversely, if a gene has an essential role in embryonic development before kidney formation and results in embryo lethality before renal development, then its function or expression during nephrogenesis will not be fully realized. Therefore, alternative methods would have to be used, such as *lacZ* tagging of the protein and analysis of the resulting transgenic mice, to determine whether the gene was expressed in the kidney.

It is possible that alternative pathways in the mouse physiology compensate for lack of a gene's function. Transforming growth factor-β1 (TGF-β1) knock-out mice are normal at birth (Shull *et al.* 1992), but this does not signify the lack of a role for the gene during kidney development. The TGF-β1 protein is expressed in the mesenchyme underlying epithelial structures in the kidney (Heine *et al.* 1987), reflecting mRNA expression (Lehnert & Akhurst 1988). *In vitro* studies have also shown that there is enhanced nephrogenesis when TGF-β1 antibodies are applied in culture, suggesting that branching morphogenesis is inhibited by the growth factor (Sakurai & Nigam 1997).

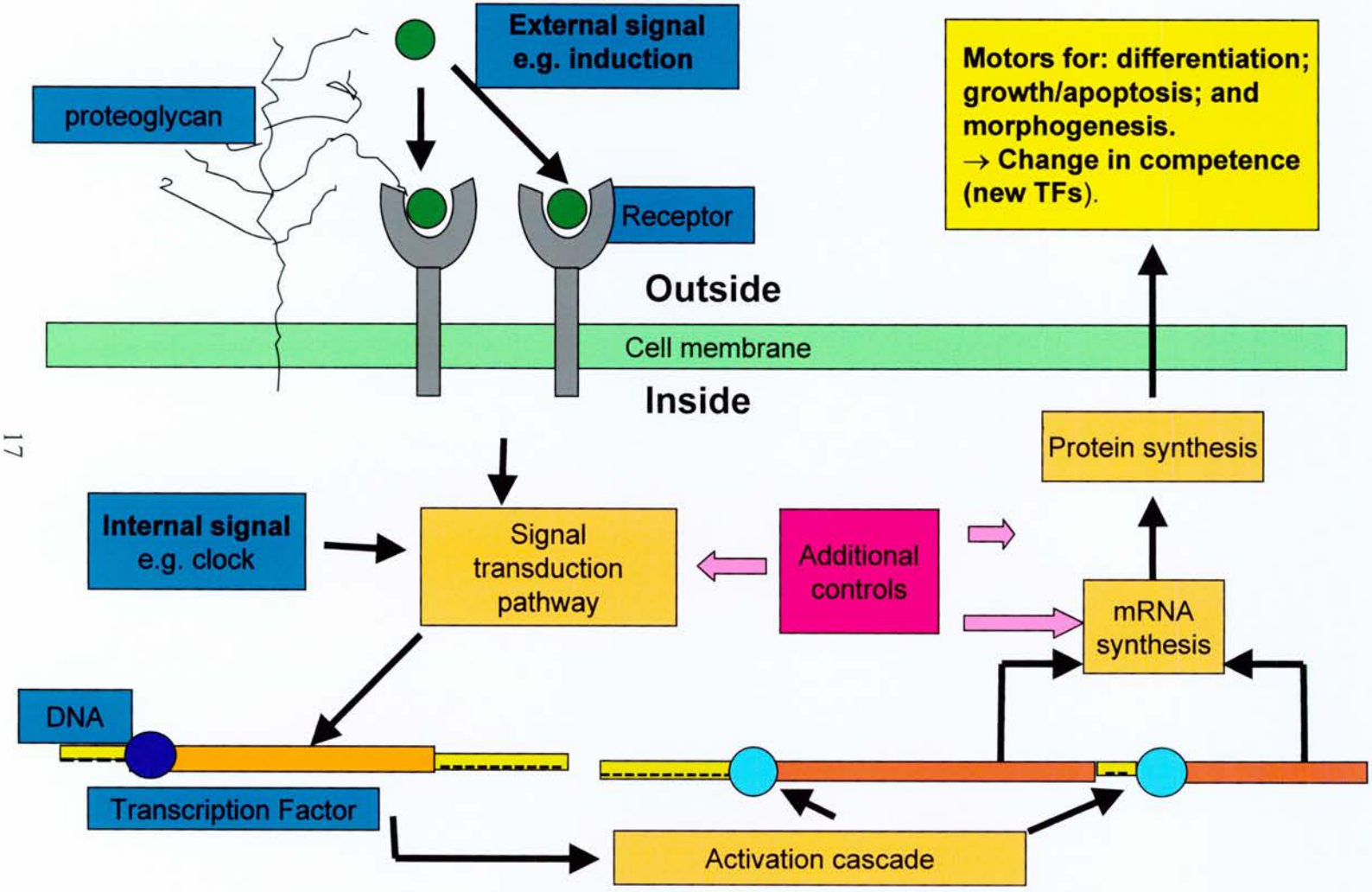


Figure 1.6. Typical genetic network for development. Courtesy of Dr Jonathan Bard.

Thus, compensatory pathways may allow for normal development after knock-out of the gene, which could lead to misinterpretation of the results since the initial conclusion would be that the gene had no function at that particular stage of development.

1.5.1. Transcription factors

Transcription factors are nuclear proteins which regulate gene expression. One of the earliest genes expressed during kidney development is the Wilm's tumour suppressor gene (*WT-1*) present in the mesenchyme before UB formation, and in the comma- and S-shaped structures after induction (Pritchard-Jones *et al.* 1990). Mutations in the *WT-1* gene can result in childhood renal tumours. *Pax-2* is a nuclear transcription factor of the paired-box family and, as mentioned earlier, is expressed by the nuclei of condensing mesenchyme and the epithelial derivatives, but the gene is down-regulated as the kidney matures (Dressler *et al.* 1990; Dressler & Douglass 1992). *Pax-2*$^{-/-}$ mice lack kidneys and even *Pax-2*$^{+/-}$ mice have kidney hypoplasia (Torres *et al.* 1995). *WT-1*$^{-/-}$ mice do not express *Pax-2* in the mesenchyme, suggesting that *WT-1* is expressed upstream of *Pax-2* during kidney development (Kreidberg *et al.* 1993).

The *Lim1* homeodomain gene encodes a transcription factor that controls early inductive events during renal development (Barnes *et al.* 1994). The protein is expressed in the UB, induced MM aggregates and the developing nephron (Fujii *et al.* 1994; Karavanov *et al.* 1998; for review, see Piscione & Rosenblum 1999). *Lim1*$^{-/-}$ mice have a lethal phenotype as well as renal and gonad agenesis, and they lack anterior head development (Shawlot & Behringer 1995).

As described earlier, the *BF-2* (*Foxd1*) gene codes for a nuclear binding protein that is expressed in the stroma of the developing kidney. The stroma is an important source of growth factors for the developing kidney, and transgenic mice lacking *BF-2* expression do not have a duct system and have few nephrons (Hatini *et al.* 1996). This transcription factor has an important role in three separate

processes: differentiation of the condensed mesenchyme, growth and branching of the UB, and elongation of the ureter.

Analysis of heterozygous and homozygous mice for other genes of the forkhead/winged helix family, *Foxc1* and *Foxc2* (formerly known as *Mf1* and *Mfh1*, respectively), has suggested a role for these overlapping genes in kidney and cardiovascular development (Kume *et al.* 2000). Compound heterozygous mice have hypoplastic kidneys and a single hydroureter, whereas the heterozygotes are normal.

1.5.2. Signalling molecules, receptors and signal transduction pathways

External factors, such as growth factors and hormones, can have a direct effect on gene transcription through the activation of cell surface receptors which then excite the intracellular signal cascade pathways. This results in transmission of the signal through the cytoplasm of the cell to the nucleus, altering gene transcription and leading to effects on morphogenesis and development (Fig. 1.6).

Growth factors are bioactive peptides which have a range of functions. These peptides act upon the cell by binding to specific cell surface receptors, causing a cascade reaction within the cell. The two main classes of growth factor receptors are the protein tyrosine kinases and the guanine nucleotide-binding (G) proteins. Growth factors may control development by autocrine (affecting the same tissue type) or paracrine (affecting a different tissue type) control pathways (for review, see Mumby & Walter 1993). Adding to the complexity of the regulatory systems involved, the factors may be regulated by more than one signalling pathway in the cell, depending upon receptor expression.

The phosphorylation and dephosphorylation of regulating proteins controls cellular processes, such as the cell cycle and protein synthesis. Protein kinases are enzymes which phosphorylate proteins (i.e. add a phosphate group to a protein) in the cell and protein phosphatases dephosphorylate proteins (i.e. remove a

phosphate group from a protein). This balance between the phosphorylated state of the target proteins is important for normal growth and development.

Protein kinases comprise of serine/threonine kinases and tyrosine kinases. The mitogen-activated protein kinases (MAPKs) are a major serine/threonine kinase group involved in signal transduction, and there are at least three MAPK signalling pathways: extra-cellular signal-regulated kinase (ERK); c-jun N-terminal kinases (JNKs); and p38-MAPK. Since regulation of the phosphorylated state of a protein affects the cell cycle, growth (and consequently development), then abnormal receptor expression, altered cytokine/hormone levels or disruption of the signal transduction pathway can have a direct effect on morphogenesis. Protein phosphatases are discussed in further detail in section 1.7.

There are many signalling molecules expressed in the developing kidney and one of the main groups is the TGF-β superfamily. TGF-β1 can inhibit nephrogenesis, although the growth factor has no effect on uninduced tissue (Rogers *et al.* 1993). Experiments using conditioned media and cell lines cultured in collagen gels have shown that TGF-β1 influences the nephrogenic epithelia by: inhibition of branching morphogenesis only; inhibition of branching morphogenesis and tubulogenesis; and inhibition of growth factors influencing morphogenesis and tubulogenesis (Sakurai & Nigam 1997). However, as discussed previously in section 1.5, *TGF-β1*$^{-/-}$ mice have no obvious renal abnormalities (Shull *et al.* 1992) a finding that conflicts with *in vitro* analyses, suggesting that there are compensatory signalling pathways allowing for normal development in the absence of this growth factor. Kallapur *et al.* (1999) showed that lethality caused by *TGF-β1* gene deletion was mouse strain-dependent and that a genetic component influences susceptibility to the mutation.

GDNF is also a member of the TGF-β superfamily and is a survival factor for dopaminergic neurons (Lin *et al.* 1993). As described earlier in sections, GDNF is expressed by MM, and is required for early induction and subsequent growth of the UB. *GDNF*$^{-/-}$ mice have kidney agenesis or dysgenesis, and lack enteric

innervation (Pichel *et al.* 1996; Pepicelli *et al.* 1997; Sainio *et al.* 1997). The GDNF tyrosine kinase receptor, c-ret, is expressed by the UB. Disruption of *c-ret* expression results in similar development abnormalities to the *GDNF* knock-out mice (Durbec *et al.* 1996; Schuchardt *et al.* 1994, 1996). In *GDNF* mutants, the MM continues to express *Pax-2* at E12·5, indicating that *GDNF* may be expressed downstream of *Pax-2* (Pichel *et al.* 1996).

As mentioned earlier in section 1.4, VEGF is an angiogenic cytokine important for blood vessel development within the kidney. Carmeliet *et al.* (1996) showed that formation of blood vessels was abnormal (but not absent) in heterozygous $VEGF^{+/-}$ embryos, but is more impaired in $VEGF^{-/-}$ homozygotes, resulting in death at mid-gestation. This suggests dose-dependent regulation of embryonic vessel development by VEGF. Similarly, Ferrara *et al.* (1996) disrupted the VEGF gene in embryonic stem cells and showed that the loss of a single VEGF allele is lethal in the mouse embryo between E11 and E12. Angiogenesis and blood-island formation are impaired, resulting in several developmental abnormalities.

There are several known ligands for epidermal growth factor receptor (EGFR), including EGF, TGF-α and heparin-binding EGF. Sakurai *et al.* (1997) demonstrated incomplete branching morphogenesis by inhibition of EGFR and HGF *in vitro*, suggesting that there are other tubulogenic pathways present in the developing kidney. The interaction between TGF-α, EGF and EGFR has been implicated in polycystic kidney disease (Fouser & Avner 1993). Richards *et al.* (1998) mutated EFGR and then introduced the gene into mice affected by autosomal recessive polycystic kidney disease (ARPKD). The kidney function improved, suggesting that EGFR has a role in the formation of kidney cysts. HGF is expressed by the MM, and transcripts of *c-met* (its receptor) have been localized to the ureter, proximal and distal tubule (Sonnenberg *et al.* 1993). Kjelsberg *et al.* (1997) showed that the HGF and c-met pathway is not vital for the mesenchyme-to-epithelial transformation in the kidney since cultured $met^{-/-}$ cells respond to EGF and TGF-α, and undergo conversion. However, the cells form tubules in the

presence of TGF-α, but not HGF. Increased levels of interleukin-6 (IL-6) and TNF-α are both associated with glomerulonephritis and tubulointerstitial nephritis. Leonard *et al.* (1999) showed that TNF-α induced IL-6 by activation of the p38-MAPK and ERK pathways in human cultured primary mesangial and proximal tubular cells.

There is growing evidence that fibroblast growth factors (FGFs) and their pathways have a role during kidney development since mice lacking the broad-range FGF receptor (FGFR) have either small kidneys or lack renal development. There are also lung, cutaneous, glandular and limb disorders associated with this FGFR disruption (Celli *et al.* 1998). A dominant-negative mutant *FGFR* disrupted early inductive signalling in the affected tissues, indicating that FGF is required for growth and patterning. Karavanova *et al.* (1996) successfully cultured metanephric mesenchyme *in vitro* in the absence of inductive tissue using cultured medium from a UB cell line supplemented with basic FGF (bFGF or FGF-2) and TGF-α. This showed that the physical contact between the MM and UB is not necessary for induction, but does depend on a range of soluble factors present in the medium. Qiao *et al.* (1999b) analysed *FGF-7*$^{-/-}$ mice, and found that the kidneys from these mice were smaller, the number of nephrons was significantly lower and the collecting system was reduced in size in comparison to the wild-type mice. The presence of FGF-7 was found to be vital for UB growth and survival, and high levels of FGF-7 delayed differentiation. Short term exposure of exogenously applied FGF-7 to kidney cultures resulted in an increase in nephron number and collecting system size.

The family of *Wnt* genes encode for secreted glycoproteins, some of which are expressed in the developing kidney. Kispert *et al.* (1996) showed that *Wnt-4* is expressed in the condensing mesenchyme, and comma- and S-shaped bodies (first shown by Stark *et al.* 1994), *Wnt-7b* is present in collecting duct epithelia, and *Wnt 11* is restricted to the branching ureteric tips. *Wnt-4* is required for tubule formation and can induce tubulogenesis in isolated mesenchyme culture, but *Wnt-11* cannot (Kispert *et al.* 1998).

It is apparent that a range of growth factors is required for normal kidney development, as highlighted by the *in vitro* experiments carried out by Karavanova *et al.* (1996) but more recently by Plisov *et al.* (2001). Plisov *et al.* (2001) has shown that cultured rat UB cells secrete TGF-β2 and leukaemia inhibitory factor (LIF), factors which resulted in enhanced tubulogenesis in isolated, uninduced rat MM cultures, although TGF-β2 can induce MM epithelialization in the absence of LIF. Furthermore, the experiments suggested that a common Wnt-dependent pathway is shared by TGF-β2, LIF and FGF-2, since exogenous Frizzled-related protein 1 (which neutralize Wnt ligands) negated the stimulatory effect of the factors.

1.5.3. Components of the extra-cellular matrix

The extra-cellular matrix (ECM) provides the necessary framework to support the branching ureteric bud and has a role in controlling the differentiation and morphogenesis of the developing kidney (Fouser & Avner 1993). The cell–cell or cell–matrix contact dictates differentiation of a cell because of the presence of cell adhesion receptors and the growth factors in the environment, resulting in excitation of one of the signal transduction pathways (for review, see Sastry & Horwitz 1996).

The basal lamina is composed of cell adhesion molecules, proteins, proteoglycans and a range of receptors, and the collagen and laminin components form sheets which are bound by nidogen (also referred to as entactin; Ekblom 1996). The basal lamina changes in composition as the kidney develops (Fouser & Avner 1993), a process that is important since the membrane must support the change in the physiology of the mesenchymal cells and regulate the cell polarity of the epithelial cells (Hay 1995). Before kidney induction, the basal lamina consists of fibronectin, and collagen types I and III, but after induction, collagen type IV and laminin predominate. Polycystic kidney disease is a result of disruption of the

basal lamina, although other elements, such as altered growth factor levels, are also involved.

Laminins are heterotrimeric glycoproteins and are one of the main components of basement membranes. *In vitro* experiments using a blocking antibody specific to the laminin E8 cell-binding site resulted in inhibition of epithelial structures (Sorokin *et al.* 1990). As described earlier, the glomerular basement membrane is an important structure of the glomerular filtration unit, but is also essential for vascular organization during development since homozygous mice lacking the gene for laminin $\alpha 5$ lack properly vascularized glomeruli. Furthermore, some of these mice have only one or no kidney, suggesting a role for $\alpha 5$ in kidney development (Miner & Li 2000).

Integrins are heterodimers consisting of α and β peptides, and are the main cell surface receptors of the basal lamina (for review, see Ekblom 1996). As the kidney develops, there is an alteration in the expression pattern of the α and β chains (Korhonen *et al.* 1990). The $\alpha 3\beta 1$ integrin receptor, expressed by the UB and derived ducts, has a role during branching morphogenesis (Ekblom *et al.* 1991; Rahilly & Fleming 1992). The $\alpha 3\beta 1^{-/-}$ mice have abnormal kidneys, exhibiting irregularities including lack of organization of the basement membrane and reduced branching of the glomerular capillary loops. The lungs also show signs of decreased branching and the mice die on the first day after birth (Kreidberg *et al.* 1996).

The $\alpha 8\beta 1$ integrin is expressed in the MM before and after condensation but not in later structures. Heterozygous mice which carry one mutated $\alpha 8$ locus are phenotypically normal. Most of the homozygous mice lacking this $\alpha 8$ subunit die soon after birth as a result of kidney agenesis or dysgenesis, although a small number of animals survive because of the presence of one kidney. These findings show the importance of integrins for kidney morphogenesis (Muller *et al.* 1997).

1.5.4. Other factors involved in kidney development

Sulphated glycosaminoglycans (GAGs) are proteins expressed by both the ureteric bud and nephrons. Syndecan-1, the first identified GAG to be expressed during kidney development, increases in expression during mesenchymal condensation. Davies *et al.* (1995) used an *in vivo* culture system to show that GAGs regulate growth and morphogenesis of the UB, and inhibition of the UB (but not nephron formation) occurs if sodium chlorate, an inhibitor of GAG sulphation, is present in the culture medium. Expression of the secreted glycoprotein *Wnt-11* is lost when metanephric kidneys are cultured in the presence of chlorate, suggesting that sulphated proteoglycans are important for sustained *Wnt-11* expression (Kispert *et al.* 1996). Kispert *et al.* (1998) also showed that *Wnt-4* signalling requires GAGs and cell contact for tubulogenesis. Heparin sulphate proteoglycans (HSPGs) interact with a range of signalling molecules, including members of the Wnt family. Homozygous mice lacking the gene encoding heparin sulphate 2-sulphotransferase (HS2ST) do not develop kidneys, have increased bone mineralization of the skeleton and show signs of reduced eye development (Bullock *et al.* 1998).

Lithium ions can induce epithelial morphogenesis of metanephric mesenchyme without the presence of an inducing tissue *in vivo* (Davies & Garrod 1995). *In vitro* experiments by Klein & Melton (1996) have shown that lithium inhibits the enzyme glycogen synthase kinase-3β (GSK-3β), thereby mimicking the Wnt signalling pathway. Whatever the precise mechanisms involved, these experiments showed that a chemical inducer of mesenchymal differentiation can replace the inducing tissue for *in vitro* experiments.

1.5.5. Summary

The complex interaction of the transcription factors, signalling molecules and ECM receptors all have a role in the control of embryonic growth and development. In terms of metanephric formation, WT-1 is one of the first factors expressed, but normal kidney development involves sequential and parallel expression of a range of genes after initiation.

1.6. Protein phosphatase inhibitor-1 may be a stem cell marker in the developing kidney

As described above, kidney-specific markers have been identified for certain cell populations (e.g. *BF-2/Foxd1* for the stromal cells; *Pax-2* is expressed by early mesenchyme-to-epithelial cells), but there is currently no unique cellular marker for the stem cells of the kidney. This sub-population of cells has an important role in forming the stromal and epithelial cells of the kidney, and it is thought that adult kidneys can partially regenerate after surgery, presumably because of the presence of these cells.

Svennilson *et al.* (1995) analysed frozen cryostat sections of E15 rat kidneys for protein phosphatase types 1 and 2A (PP-1 and PP-2A, respectively) using RNA *in situ* hybridization. Protein phosphatases are required for regulation of growth and differentiation, and Svennilson *et al.* (1995) showed that both types of phosphatases are ubiquitously expressed in most cellular compartments of the rat kidney. However, sections incubated with radio-labelled oligonucleotides for PP-1 inhibitor-1 (inhibitor-1 or I-1, an endogenous inhibitor to PP-1) show specific expression in mesenchymal cells at the periphery of the tissue. On the basis of these results, Svennilson *et al.* (1995) suggested that I-1 is the first marker specific to the cortical mesenchymal stem cells of the kidney and classified these stem cells as phenotypically unaltered by the ureteric bud. Bard *et al.* (1996) and Davies & Brandli (1997) both referred to stem cells as already induced metanephric mesenchyme (hence these cells would be expected to express *Pax-2*). The confusion surrounding the definition of a stem cell highlights the need for a suitable marker for defining such cells in the kidney. However, for the I-1 protein to be a candidate stem cell marker, it would have to fulfil certain criteria. The protein should only be expressed in a sub-population of MM kidney cells in the periphery of the organ and should co-express with other stem cell markers (e.g. Pax-2 and IIGF, although both proteins are not exclusively expressed by the stem cells). Stem cells should exhibit reduced apoptosis since these cells must survive to ensure continued nephron formation during development.

1.7. Protein phosphatases and inhibitor-1

Inhibitor-1 is a small heat- and acid-stable protein, and an endogenous inhibitor to PP-1. Protein phosphatase-1 and PP-2 are the two main classes of protein phosphatases and their function is to remove phosphate groups from a protein. PP-1 is comprised of four isoforms derived from three genes. PP-2 has three subgroups: 2A, 2B and 2C. PP-2B (or calcineurin) is more substrate-specific than PP-1 or PP-2A. There are other phosphatases (e.g. PP-6), but less is known about their tissue specificity and functions in the cell. Protein phosphatase-1 and -2A comprise over 90% of the protein/serine phosphatase activity in mammalian cells, and dephosphorylate proteins in a range of processes and pathways, including the cell cycle and protein synthesis. These phosphatases are the main antagonists to the protein kinases which phosphorylate proteins in the cell. It is this balance of phosphorylation and dephosphorylation that regulates cellular functions and processes such as protein synthesis, the cell cycle and carbohydrate metabolism (for reviews, see Mumby & Walter 1993; Oliver & Shenolikar 1998).

Phosphorylation of I-1 by cAMP-dependent protein kinase results in the suppression of PP-1 activity, leading to a large increase in intracellular phosphorylation reactions (for reviews, see Mumby & Walter 1993; Oliver & Shenolikar 1998). The unique conformational change of I-1 upon phosphorylation appears to be selectively recognized by PP-1 and I-1 is likely to have no other role in the cell than to inhibit PP-1 (Endo *et al.* 1996). Since the phosphorylation of I-1 can be altered in response to hormones and neurotransmitters, I-1 has a role in signal transduction.

The rabbit I-1 protein sequence was first derived from skeletal muscle protein by Aitken *et al.* (1982) and has strong homology with the I-1 sequences derived from: rabbit liver protein (MacDougall *et al.* 1989); the predicted protein sequence from rat muscle cDNA by Elbrecht *et al.* (1990); and the predicted human sequence from a cDNA brain library (Endo *et al.* 1996). Allen *et al.* (2000) created *I-1*$^{-/-}$ mice which underwent normal physical development. Thus, loss of the I-1 protein is not lethal to the developing embryo, but long-term potentiation (LTP) was

down-regulated at perforant path-dentate granule cell synapses, although the mice performed as well as wild types during water maze tests. LTP is important for maintaining long-term synaptic strength.

Aitken *et al.* (1982) determined that rabbit I-1 protein had a molecular weight (M_r) of 18 640 but a M_r of 26 000 was estimated after polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulphate (SDS-PAGE). The latter figure is in agreement with the M_r of 26 500 detected by Hemmings *et al.* (1992) from rabbit brain tissue and the M_r of 26 000 found by MacDougall *et al.* (1989) from a range of rabbit tissue. In mouse, the I-1 protein has an apparent molecular mass of 30·5 kDa (extracted from brain by Hemmings *et al.* 1992), and it is 30 kDa for rat (MacDougall *et al.* 1989; Hemmings *et al.* 1992; Fryckstedt *et al.* 1993; Lowenstein *et al.* 1995; all determined by SDS-PAGE). The apparently higher than expected values obtained by SDS-PAGE were a result of the abnormally low binding of the detergent (Aitken *et al.* 1982).

The expression pattern of I-1 has mainly been demonstrated by immunoblotting and Northern analysis. MacDougall *et al.* (1989) detected I-1 by immunoblotting in rabbit and rat skeletal muscle, brain, heart, kidney, uterus and adipose tissue. Inhibitor-1 was not detected in rat or mouse liver, but was present in the tissue from rabbit, pig, sheep and guinea pig. MacDougall *et al.* (1989) also detected I-1 in the skeletal muscle of a range of species, including mouse, pig, rabbit, rat and guinea pig. Hemmings *et al.* (1992) detected I-1 by immunoblotting in various compartments of the brain of frog, turtle, canary, pigeon, mouse, rabbit, cow and monkey, but not goldfish. Peripheral tissues such as heart, kidney and pancreas from a range of species (including rabbit, rat and cow) also expressed I-1 at various levels. There was no I-1 detected in rat and cow liver, a similar finding to the results obtained by MacDougall *et al.* (1989). Fryckstedt *et al.* (1993) detected I-1 in rat kidneys by immunoblotting and immunocytochemistry. Sakagami *et al.* (1994) used *in situ* hybridization and immunoblotting to detect I-1 in rat brains and showed high levels in certain cortical compartments, including the hippocampal formation and the piriform cortex. Lowenstein *et al.* (1995) showed the presence

of I-1 by immunoblotting and immunocytochemistry in the neocortex of adult rat, cat and ferret.

1.8. Functions of inhibitor-1

The role of I-1 appears to be solely to inhibit PP-1, and thus, it would be reasonable to assume that any function of I-1 will ultimately involve PP-1. The four identified areas of I-1 function are: synaptic plasticity; glycogen metabolism; cell growth; and muscle contraction (Oliver & Shenolikar 1998).

The I-1 knock-out mice developed by Allen *et al.* (2000) showed no obvious physical abnormalities. However, this does not necessarily imply lack of physiological roles in the cell for I-1 since it is feasible that alternative compensatory pathways could be involved. For example, the presence of another PP-1 inhibitor could increase the phosphorylation reactions in the cell. As mentioned earlier, the I-1 knock-out mice lacked long-term potentiation, although the animals performed equally as well as wild types during memory tests.

Adrenalin activates protein kinases in the cell and inactivates the enzyme glycogen synthase via phosphorylation. This is thought to be because of the kinase phosphorylation of I-1 leading to PP-1 inactivation (for review, see Oliver & Shenolikar 1998). Scrimgeour *et al.* (1999) compared wild type and I-1 knock-out mice to test the hypothesis that activation of glycogen synthase by insulin is caused by the lowering of PP-1 inhibition by I-1. However, there was no difference between mouse types since insulin increased glycogen synthase activity, and subsequently, glycogen synthesis in both cases. This finding suggests that I-1 is not essential in the stimulation of glycogen synthesis and the current model for glycogen synthase in cultured cells involving a PP-1 targetting subunit may not be valid. However, the work of Scrimgeour *et al.* (1999) does not exclude the possibility that there may be subsidiary pathways involved in glycogen synthesis.

Protein phosphatase-1 is necessary for the growth of cultured pituitary tumour cells. The addition of cAMP and cyclosporin A (both are antiproliferative) increased the levels of I-1 phosphorylation, and thus, I-1 may be a growth regulator (for review, see Oliver & Shenolikar 1998). However, PP-1 levels do not fluctuate during the cell cycle.

Gupta *et al.* (1996) showed that I-1 is present in guinea pig ventricular cardiomyocytes and that levels are hormonally regulated since treatment of I-1-enriched extract with isoproterenol (a β-adrenergic agonist) resulted in phosphorylation. This may be important for cardiac contractility through inhibition of PP-1.

Smooth muscle PP-1 is bound to myosin and dephosphorylation of the myosin light chain protein results in relaxation of the muscle, but an increase in intracellular levels of cAMP also leads to the relaxation of smooth muscle. This conflicting observation is a result of hormonal activation of I-1 (e.g. by epinephrine), inhibiting the PP-1 that interacts with smooth muscle myosin light chain kinase (MLCK) phosphorylation. Thus, the balance of phosphorylation and dephosphorylation is required for normal muscle control (for review, see Oliver & Shenolikar 1998).

1.9. Other inhibitors of protein phosphatase-1

Inhibitor-1 is not the only inhibitor to PP-1. Other inhibitors have different tissue distribution and different modes of PP-1 inhibition.

1.9.1. DARPP-32

Dopamine- and AMP-regulated phosphoprotein with an apparent M_r of 32 000 (DARPP-32) is structurally related to I-1, is also a cAMP-regulated inhibitor of PP-1, but is the product of a separate gene and has different tissue distribution (Hemmings *et al.* 1992; Fryckstsedt *et al.* 1993; Lowenstein *et al.* 1995).

Furthermore, the two proteins may interact differently with the PP-1 catalytic sub-unit (Connor *et al.* 1998). DARPP-32 inhibits PP1 activity only after phosphorylation on threonine 34 by protein kinase A (Walaas *et al.* 1983). DARPP-32 levels are reduced in the striatum of transgenic mice presymptomatic for Huntington's disease (Bibb *et al.* 2000).

DARPP-32 is found predominantly in the brain (particularly in the basal ganglia) and neuronal tissue in a range of species (Hemmings *et al.* 1992), and was initially extracted from bovine adipose tissue (Stralfors *et al.* 1989). However, adipose tissue from rabbit and rat mainly contains I-1 (Hemmings 1992) and, although the significance of this difference is unclear, it suggests there may be a role for both inhibitors in adipogenesis.

DARPP-32 is present in the kidney (Aperia *et al.* 1987) and the protein is concentrated more in the renal medulla than the cortex, and is present in the UB (Fryckstedt *et al.* 1993). In renal tubules, DARPP-32 inhibits sodium resorption by blocking the activity of Na^+, K^+-ATPase (Aperia *et al.* 1987; Meister *et al.* 1989).

Double knock-out DARPP-32 and I-1 mice (i.e. mice with both DARPP-32 and I-1 genes disrupted) have been generated and the regulation of glycogen synthase was found to be similar to the wild type mice (Scrimgeour *et al.* 1999).

1.9.2. Inhibitor-2

Inhibitor-2 (I-2) is an acid- and heat-stable protein (for review, see Oliver & Shenolikar 1998). In human fibroblasts, I-2 reaches its highest expression during S phase and mitosis during the cell cycle, as determined by Kakinoki *et al.* (1997) using green fluorescent protein fused to I-2. Unlike I-1, I-2 can inhibit PP-1 when unphosphorylated, and Connor *et al.* (2000) characterized a novel interaction between the N terminus of the I-2 protein and the PP-1 domain, consisting of amino acids Glu-53, Glu-55, Asp-165, Glu-166 and Lys-167. However, I-2 and DARPP-32 have separate unrelated motifs required for binding PP-1 (Huang *et al.*

1999). Osawa *et al.* (1996) isolated three separate isotypes of I-2 from a rat testis cDNA library: I-2α1, I-2α2 and I-2β. Inhibitor-2α1 and I-2α2 are derived from the same gene, but I-2β is coded from a different gene, is only expressed in the testis, and may have a role in sperm motility (Osawa *et al.* 1996).

1.9.3. NIPP-1

Nuclear inhibitor of PP1 (NIPP-1) was originally isolated from bovine thymus nuclei (Beullens *et al.* 1992) and the protein is expressed in a range of mammalian tissues (Van Eynde *et al.* 1995). NIPP-1 is similar to I-2 in that it is a potent phosphatase inhibitor in the dephosphorylated state and forms a stable, inactive complex with PP1. Wera *et al.* (1997) showed that NIPP-1 expression in reticulocyte lysates and COS-1 cells is controlled at the translation level, and thus, over-expression of this PP-1 inhibitor in the COS-1 cells was unsuccessful (NIPP-1 is one thousand times more potent as an inhibitor of PP-1 than I-1 or I-2).

1.9.4. RIPP-1

Buellens *et al.*(1996) isolated a ribosomal 23 kDa basic polypeptide inhibitor (RIPP-1) from rat liver that was more efficient in inhibiting ribosome-associated PP-1 (PP-1R) than other protein phosphatases.

1.9.5. CPI17

C-kinase activated phosphatase inhibitor (CPI17) is a novel inhibitor of PP-1. It was isolated from porcine smooth muscle (Eto *at al.* 1995) and is the same protein encoded by a cDNA clone isolated from a porcine aorta cDNA library (Eto *et al.* 1997). The CPI17 protein is 16·7 kDa and is only expressed in smooth muscle: the mRNA for CPI17 is expressed exclusively in smooth muscle such as aorta and bladder, but not in skeletal muscle or non-muscle tissues (determined by Northern analysis). CPI17 inhibits the myosin-bound PP1 complex as well as the free PP1 catalytic sub-unit, suggesting that it is functionally different to I-1 and DARPP-32. CPI17 has a role in the control of smooth muscle, and myofibril-bound PP-1 has

been identified as a target of CPI17 in vascular smooth muscle by Senba *et al.* (1999), who analysed the myofibrillar extract from porcine aorta.

1.9.6. Inhibitor-4

Shirato *et al.* (2000) isolated I-4 from a human cDNA library of germ cell tumours, and the protein shares 44% homology to I-2. Inhibitor-4 strongly interacts with PP-1C.

1.9.7. Summary

The inhibitors DARRP-32, I-2 and NIPP-1 may be present in the developing kidney but RIPP-1 and CPI17 are not likely to be found outwith their own specified tissue. However, the full array of inhibitors to protein phosphatases is not complete since the use of more sensitive molecular techniques is revealing the presence of other inhibitors. Thus, the presence of other inhibitors in the cell cannot be discounted.

1.10. The aim of the project

The initial aim of this research project was to test the hypothesis of Svennilson *et al.* (1995) that the I-1 protein is a kidney stem cell marker. Kidney stem cells differentiate to give both nephrogenic and stromal cells following induction, as described earlier, and signal transduction will be important in this process. Therefore, I-1 may have an important role in modulation of differentiation of these cells. Determining a unique cellular marker for the stem cells would have benefits in confirming the location and physiological role of these cells during kidney development. Initially, genomic clones were isolated after a mouse library screening and were analysed to ensure homology to the known rat I-1 mRNA sequence. From these clones, the genomic structure of the mouse I-1 gene was determined, the mRNA and protein sequences were predicted, and the chromosomal localization of the gene was identified (Chapter 3). The protein

expression of I-1 in the developing mouse embryo was determined using whole-mount tissue, a commercially available antibody and confocal microscopy (Chapter 4). Finally, the presence of regulatory elements in the 5' flanking sequence of the I-1 gene was analysed using reporter plasmids and transfection experiments (Chapter 5).

Chapter 2
Materials and methods

2.1. Introduction

Before any practical work was initiated, safety procedures were followed according to the Department of Biomedical Sciences safety manual (http://www.bms.ed.ac.uk/healthsafe/home.htm) and the University of Edinburgh Health and Safety Policy (http://www.safety.ed.ac.uk/h&spol). Gloves and a lab coat were worn at all times in the laboratory. If particularly hazardous substances were handled (e.g. radio-isotopes) then double gloves were worn. This provided not only personal safety but also a safe working environment for others in the laboratory.

Laboratory waste (i.e. hazardous, biological or general refuse) was disposed of in accordance with departmental and university rules. Risk assessments and forms for the control of substances hazardous to health (COSHH) were completed for chemicals, equipment and protocols used in the laboratory. The codes of practice for genetic modification and the handling of radioisotopes were followed, including the attendance of courses associated with the relevant topics. Manual instructions for equipment were read before use and all electrical equipment had been safety-tested by the university.

2.2. General materials and solutions

The general solutions used in the laboratory were used according to protocols from the laboratory manuals by Sambrook *et al.* (1989) and Ausubel *et al.* (2000), or from datasheets provided by the manufacturers. Sigma water (Sigma, St Louis, MO, USA) was used for the elution of plasmid DNA and for some of the solutions, since it is ultra-pure.

Antibiotics Ampicillin (disodium salt) and kanamycin were dissolved in Sigma water to a final concentration of 50 mg mL^{-1} and 25 mg mL^{-1}, respectively, and stored at –20 °C.

L-Broth Miller's modification medium (Luria broth base; Sigma) was used at 15·5 g L^{-1}. The solution was autoclaved and stored at room temperature.

LB agar plates L-Broth was prepared as above and Select agar (Sigma) was added at 15 g L^{-1}, autoclaved then cooled to 55 °C in a waterbath. The plates were poured in the laminar flow hood, allowed to dry and stored at 4 °C until required. For antibiotic plates, L-B agar was prepared and after cooling to 50 °C, antibiotic was then added before pouring the plates: ampicillin and kanamycin were used at final concentrations of 50 μg mL^{-1} and 25 μg mL^{-1}, respectively. The plates were stored at 4 °C until required.

PBS A PBS tablet (Sigma) was dissolved in 200 mL distilled water, then autoclaved. This gave a final concentration of 10 mM phosphate buffer, 2·7 mM potassium chloride, 137 mM sodium chloride.

4% PFA Four grams of paraformaldehyde (PFA; Sigma) per 100 mL PBS was heated to 90 °C to dissolve, then stored at –20 °C.

TBE 5 × TBE was prepared by dissolving one sachet (Sigma) in one litre of deionized water.

TBS 100 mM Tris-HCl pH 7·5

400 mM sodium chloride

The solution was autoclaved before use.

TE Buffer 10 mM Tris

1 mM EDTA

The solution was pH'd to 7·5 then autoclaved.

SM 0·1 M Sodium chloride

10 mM magnesium sulphate $MgSO_4.7H_2O$

50 mM Tris.HCl, pH 7·5

The solution was autoclaved then gelatin was added, to give a final concentration of 0·01%.

SOB 2% w/v Bacto-tryptone (Sigma)

0·5% w/v Bacto-yeast extract (Sigma)

10 mM sodium chloride

2·5 mM potassium chloride

The solution was autoclaved and sterile 1 M magnesium chloride was added to give a final concentration of 20 mM.

SOC A 1 M glucose solution was added to SOB (made up as above) to give a final concentration of 20 mM.

SSC (20 ×) 3 M sodium chloride

0·3 M Tri-sodium citrate

The solution was diluted to give a 1 × stock as required.

Xgal/IPTG Plates A 10% w/v solution of Xgal in Dimethyl Formamide (DMF) was prepared. IPTG was dissolved to a concentration of 0·2 M in Sigma water. 20 µL each of Xgal and IPTG were spread over a dried L-agar plate. The plates were allowed to dry in the fume hood for 30 min before plating the cells.

2.3. General methods

2.3.1. Growth conditions for bacterial cultures

A liquid culture of bacterial cells was grown overnight in a 37 °C shaking incubator at 225 rotations per minute (rpm). Streaked-out cultures were grown inverted on L-agar plates in a 37 °C incubator. The plates or liquid cultures were not allowed to grow for over 17 h, to ensure optimal growth conditions for the bacterial cells and to maintain maximum plasmid integrity. The genetic manipulation room in the Anatomy Section (Department of Biomedical Sciences) was level containment 1 and was sufficient for all genetic modification work carried out for the project.

2.3.1.1. Plates: A streaked-out plate was obtained by initially sealing the end of a glass Pasteur pipette. Bacteria (e.g. from a frozen glycerol stock or a scraping from an older streaked-out plate) were transferred using the closed end of the pipette and then streaked onto a fresh plate, according to the method in the laboratory manual by

Sambrook *et al.* (1989). For plating a liquid culture (e.g. transformation of competent cells), 100 µL–200 µL of culture was spread onto a fresh plate using a sterile glass spreader. After overnight growth, the plates were sealed with Parafilm and stored inverted at 4 °C until required.

2.3.1.2. Liquid: A sterile Gilson yellow tip was used to inoculate 5 mL of L-broth (in a 30-mL sterile plastic Universal tube) using a single, well isolated colony from a streaked-out plate. The cap was loosely taped to allow adequate aeration of the culture. A sterile 50-mL plastic Falcon tube was used if the volume of culture was between 5–15 mL. Larger volumes of culture (i.e. more than 15 mL) were grown in autoclaved, sterile conical flasks.

After overnight growth of the liquid culture, a 15% glycerol stock was prepared as follows: 150 µL glycerol (sterilized by autoclaving) was added to 850 µL liquid culture in a sterile 1·5-mL Eppendorf tube. The contents were vortexed to mix and stored at –70 °C. This glycerol stock was used to re-streak the culture to avoid re-transformation of the plasmid DNA. Thus, only one clone was used for analysis.

2.3.2. Preparation of competent *Escherichia coli* strain DH5α cells

Competent cells *Escherichia coli* (*E. coli*) strain DH5α were prepared using the calcium chloride protocol described in the laboratory manual by Ausubel *et al.* (2000).

2.3.3. Transformation of competent *Escherichia coli* strain DH5α cells

(1) An aliquot of competent cells was removed from the –70 °C freezer and thawed on ice. Fifty-mL or 30-mL sterile plastic tubes were cooled on ice.

(2) A 20 µL aliquot of the competent cells was placed into each cooled tube and 1 µL of plasmid DNA was added. The contents were swirled gently to mix, and the tubes incubated on ice for 30 min.

(3) The cells were heat-shocked for 40 s in a 42 °C waterbath and returned to ice for 2 min. Eighty microlitres of SOC was added to the cells and the tubes were incubated at 37 °C, 225 rpm, for 1 h.

(4) The cells were plated onto L-agar plates with ampicillin or kanamycin (used at 50 and 25 μg mL^{-1}, respectively), as described above in section 2.3.1. The plates were grown inverted overnight at 37 °C.

(5) After growth of the colonies, a sterile yellow tip was used to inoculate 5 mL L-Broth plus antibiotic. The cultures were grown overnight in the 37 °C shaking incubator at 225 rpm.

When scaling up the transformation reagents, 100 μL competent cells and 5 μL plasmid DNA were used but the cells were heat shocked for 50 s then added to 100 μL of warmed SOC medium.

2.3.4. Transformation of XL10-Gold ultra-competent cells

This procedure was carried out according to the instructions supplied by Stratagene (La Jolla, CA, USA).

2.3.5. Plasmid DNA preparation

Figure A1a–g, Appendix 1 (pp. 166–172) lists the maps of the vectors used in this research project. QIAprep Spin miniprep kit (QIAGEN GmbH, Hilden, Germany), PERFECT prep miniprep kit (5 Prime → 3 Prime, Inc., Boulder, CO, USA), GFX™ Micro plasmid prep kit (Amersham Pharmacia Biotech, Little Chalfont, UK), Wizard® *Plus* MiniPrep and Wizard® *Plus* MaxiPrep DNA Purification System (Promega Corporation, Madison, WI, USA) were all used according to the manufacturer's instructions. All plasmid miniprep DNA was eluted in a final volume of 100 μL. However, if the insert size was larger than 10 kb, the Sigma water was heated to 70 °C to before elution, as recommended by the QIAGEN handbook. This step was included for each miniprep kit where appropriate. The maxiprep DNA was eluted in 1 mL of Sigma water.

2.3.6. Restriction digestion of DNA

Reaction volumes for digestion of DNA depended upon the concentration of DNA used. The values given in Table 2.1 were a guide for setting up single, double and triple digests, assuming that the DNA concentration was between 500 and 1000 ng μL^{-1}. If more DNA was required for digestion (e.g. for excising a band from a low melting point agarose gel), then the reaction volumes were increased, but the volume of reaction buffer was maintained at 10% of the final volume. Routinely, the volume of enzyme used was 1 µL of stock. Most enzymes had a concentration of 10 units μL^{-1}, and since 1 unit of enzyme is defined as digesting 1 µg of DNA in one hour at 37 °C, then adding excess enzyme should allow for complete digestion.

Table 2.1. The volume of reagents (µL) for restriction digestion of DNA. The volume of water used per digest was dependent on the concentration of DNA, and thus, values 'x' and 'y' were variable.

Reagent	Single digest	Double digest	Triple digest
DNA	x	x	x
10 × Buffer	1	1	1
Enzyme 1	1	1	1
Enzyme 2	–	1	1
Enzyme 3	–	–	1
Sigma water	y	y	y
Final volume	10	10	10

2.3.7. Agarose gel electrophoresis

Ethidium bromide was handled carefully, and contaminated solid waste was disposed of according to local health and safety rules. When viewing a gel using the ultra violet (UV) transilluminator, gloves and a full-face mask were worn. Any liquid waste containing ethidium bromide was passed through an ethidium bromide adsorption column (BDH, Poole UK) to remove the chemical. The pass-through liquid was rendered safe by the column.

The electrophoresis buffer, TAE, was prepared according to the method described in the laboratory manual by Sambrook *et al.* (1989) and TBE was prepared as described in section 2.2. Agarose from Sigma (type EEO) was prepared in 1 × TBE for standard gel electrophoresis. SeaKem agarose (BioWhittaker Molecular Applications, Rockland, ME, USA) and low-melting-point agarose (Sigma) were prepared in 1 × TAE for optimal recovery of DNA from the gel. Markers were included in the gels to allow for both prediction of the size of bands and also to estimate the concentration of DNA. The 1-kb ladder from Life Technologies Inc. (GIBCO BRL, Rockville, MD, USA) was used routinely.

(1) Powdered agarose was added to 1 × TBE or 1 × TAE in a conical flask, and loosely covered with cling film. The solution was weighed and the mass recorded. The agarose solution was carefully heated in the microwave until the agarose was dissolved, but care was taken to ensure that the solution did not over-boil. The weight of the flask was noted and distilled water was added if there was a loss of mass (otherwise a more concentrated agarose concentration would result). The agarose solution was placed in a 55 °C waterbath to cool before pouring into the mould.
(2) The gel apparatus was set up according to the manufacturers instructions. The gel moulds were cleaned in 70% ethanol and the ends taped using insulating tape.
(3) Ethidium bromide (Sigma) was added to a final concentration of 125 ng mL^{-1} to the cooled agarose before casting. The gel was allowed to set at room temperature.
(4) Once the agarose had solidified, the tape and comb were removed from the gel mould and the assembly placed into the tank of 1 × TBE or 1 × TAE.
(5) DNA loading buffer (Sigma) was added to samples of digests at a dilution of approximately 1:5 with sample. An appropriate DNA ladder was also prepared with gel loading buffer. The samples were then loaded into the wells of the gel.
(6) The gel tank was then connected to a power supply (e.g. midi sized gels, were set at 25 V overnight).

(7) Once the gel front had migrated a suitable distance along the gel, the gel was removed and photographed using a Polaroid land camera and UV transilluminator.

Midi-sized gels were often used, but ethidium bromide cannot be added directly to the melted agarose gel before casting. After electrophoresis overnight, the gel was stained in a solution of 250 ng mL^{-1} ethidium bromide (diluted from a stock of 500 μg mL^{-1}).

2.3.8. GENECLEAN protocol

The GENECLEAN kits were used to purify DNA according to the manufacturer's instructions (BIO 101, Vista, CA, USA). The GENECLEAN and GENECLEAN Spin kits were used.

2.3.9. Ligation reactions

Ligation of DNA into Promega pGEM®-T Easy vector (Fig. A1a, Appendix 1, p. 166) was carried out according to manufacturers instructions (see Table 2.2 for volumes used). The vector, 10 × buffer and T4 DNA ligase were provided in the kit. The tubes were either incubated at 22°C for several hours or overnight at 4°C.

Table 2.2. The volume of reagents (μL) used for ligation of PCR product into vector pGEM®-T Easy (Promega).

Reagent	Normal	Control
336-bp PCR DNA	1	–
10 × Ligase Buffer	1	1
pGEM®-T Easy Vector	1	1
T4 DNA Ligase	1	1
Sigma water	6	7
Final volume	10	10

Ligation of insert DNA into vector will only be successful if the insert and vector DNA strands have compatible ends. Complementary overhanging ends are more efficient at ligation than blunt-ended DNA. However, if the vector has identical ends (i.e. if the DNA was digested using one enzyme only), then there is the potential of self-ligation, resulting in no ligation of the insert DNA. Treating the vector with calf intestinal alkaline phosphatase (CIAP) reduces this risk (Doyle 1996). Therefore, after digestion with the appropriate enzyme, one unit of CIAP (Roche Diagnostics Ltd, Lewes, UK) was added to the digest, incubated for 15 min at 37 °C and then heated to 65 °C to remove any CIAP activity. The reaction mixture was then GENECLEANed. If the vector had been digested with two different enzymes which had incompatible ends, the DNA was GENECLEANed without CIAP treatment since self-ligation was not likely to occur.

The concentration of insert used per ligation reaction was calculated using the following equation from Doyle (1996):

$$\frac{\text{ng of vector} \times \text{kb insert size}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng insert}$$

The ligation reactions were set up with the desired molar ratio of insert:vector at 3:1 (Table 2.3). Therefore:

$$\frac{\text{ng of vector} \times \text{kb insert size}}{\text{kb size of vector}} \times \frac{3}{1} = \text{ng insert}$$

Table 2.3. The volume of reagents (μL) used for ligation of insert into vector. Values 'w', 'x' and 'y' were variable.

Reagent	Control	Ligation
Insert DNA	w	w
Vector DNA	x	–
10 × Ligase buffer (Roche)	1	1
T4 DNA Ligase (Roche)	1	1
Sigma water	y	z
Final volume	10	10

A control ligation reaction was included (i.e. no insert DNA present) to determine if self-ligation had occurred. The final reaction volume was usually 10 μL or 20 μL but the quantity of 10 × buffer was increased to 2 μL if using the larger volume. The quantity of total quantity of DNA per ligation reaction was usually 50 ng.

2.3.10. Polymerase chain reaction (PCR)

Ready-To-Go PCR beads (Amersham Pharmacia Biotech) were used according to the manufacturer's instructions. The Hybaid PCR machine (ThermoHybaid, Ashford, UK) was programmed for 40 cycles, each cycle consisting of denaturation for 30 s at 95 °C, annealing of primers to template at 55 °C for 1 min, and finally, extension of the primer to give amplification of template at 72 °C for 1 min. Routinely, one-fifth of the reaction mixture was analysed by agarose gel electrophoresis.

2.3.11. Southern blot

The protocol described in the laboratory manual by Ausubel *et al.* (2000) was followed.

2.3.12. Radioactive labelling of DNA and hybridization of filters

All radioactive work was carried out in a designated room to contain the radioisotopes. The work area was monitored for contamination before and after use. The High Prime kit (Roche) was used to label the DNA, according to manufacturer's instructions.

Church Hybridization Mix 0·5 M PO_4 pH 7·2 (made from 1M $NaPO_4$ stock)
7% SDS
1 mM EDTA
1% BSA (bovine serum albumin)

Church Wash 1% SDS
1 mM EDTA
40 mM $NaPO_4$ pH 7·2

The Southern blot filters were wetted with 2 × SSC and placed in Hybaid hybridization tubes. If several filters were being hybridized, a sheet of nylon membrane was sandwiched between the filters to allow adequate flow of the radio-labelled probe. Approximately 40 mL of prehybridization solution was used per bottle. Just before use, denatured salmon sperm DNA was added to a final concentration of 100 μg mL^{-1}. The tubes were prehybridized at 65 °C in the Hybaid oven for 60 min to equilibrate the membrane. The prehybridization solution was discarded and replaced with 20 mL of fresh hybridization mix and salmon sperm. The ^{32}P-radio-labelled DNA was added and the filters were hybridized overnight at 65 °C.

The hybridization mix was carefully discarded into the drainage system using plenty of running water. The waste was monitored to prevent both splashes and the formation of an aerosol. The filters were washed at 65 °C in the hybridization oven for 3 × 20 min. The filters were then carefully wrapped in cling film and taped into X-ray cassettes and exposed onto photographic film (Fuji or Kodak). After exposure at –70 °C or room temperature (depending upon the strength of signal), the film was developed.

2.4. Specialized materials for Chapter 3

LB Broth Miller's modification medium (Sigma) and sodium chloride were added at 15·5 and 4·5 g L^{-1} of water, respectively. The solution was autoclaved and stored at room temperature.

LB agar plates LB Broth was prepared as described above and Select agar (Sigma) was added at 15 g L^{-1}. The solution was autoclaved then cooled in a 50 °C waterbath. The plates were poured in the laminar flow hood, allowed to dry and then stored at 4 °C.

Antibiotic plates LB agar was prepared as described above but antibiotics were added before pouring the plates.

LB/MgSO4 LB agar was prepared as described above, but sterile 1 M MgSO4 was added to a final concentration of 10 mM before pouring the plates.

LB++ LB broth was prepared as above then autoclaved. One molar $MgSO_4$ was added to a final concentration of 10 mM, and maltose to a final concentration of 0·2% (from a stock of 20%). The solution was stored at room temperature.

20% Maltose Twenty grams of maltose was dissolved in 100 mL sterile water and then filter sterilized using a Nalgene 0·22 μm filter unit.

LB/Top agarose LB broth was prepared as above and added $MgSO_4$ to give a final concentration of 10 mM (from a stock of 1 M), then added agarose to 7·2 g L^{-1}. The solution was autoclaved and stored at room temperature until required. Before use, the LB/top agarose was liquified in the microwave or in a boiling waterbath and cooled to 50 °C.

2.5. Specialized methods used in Chapter 3

2.5.1. I.M.A.G.E. clone analysis

In order to obtain cDNA clones corresponding to the mouse I-1 mRNA sequence, the dbEST database was screened using a keyword search. The clones were ordered through the HGMP-RC (UK Human Genome Mapping Project Resource Centre, Hinxton, Cambridge, UK), an authorized distributor for the I.M.A.G.E. consortium. Only the first few hundred base pairs of the clones had been sequenced by I.M.A.G.E., and therefore, there was only partial homology to the deduced rat I-1 mRNA.

2.5.2. Isolation of the 5' PCR product and 3' restriction fragment of 1912-d05 into vectors

The primers used for isolation of the PCR product were synthesized by MWG-Biotech UK Ltd (Milton Keynes, UK). The PCR product was amplified using clone 1912-d05 as template and Ready-To-Go beads from Amersham Pharmacia Biotech (as described in section 2.3.10). The PCR product was excised from a low-melting-point agarose TAE gel, GENECLEANed and then inserted into pGEM®-T Easy vector. Plasmid DNA was isolated after transformation of DH5α competent cells.

A 3' restriction fragment of clone 1912-d05 was excised from a low-melting-point agarose gel (Sigma) then ligated into pBluescriptIIKS–. Plasmid DNA was isolated after transformation of DH5α competent cells.

2.5.3. Library plating

The MoBiTec λPS library (male, 129 SV D3; Fig. A1b, Appendix 1, p. 167) and the plating cells (C600 and BNN132) were handled according to the information booklet provided by the manufacturer (NBL Gene Sciences Limited, Cramlington, UK). Likewise, all media and solutions were prepared according to the MoBiTec protocol. The titre of the library was initially established to allow representation of all of the mouse genome before the screening could commence.

2.5.4. Screening of the library

2.5.4.1. Primary screening: Reinforced cellulose nitrate membrane (Optitran BA-S 83, 0·2 μm, 200 mm × 200 mm) was used according to the manufacturers instructions (Schleicher and Schuel, Dassel, Germany). The first lifts (labelled 1.1 to 10.1) and the second lifts (1.2 to 10.2) of the library were hybridized with the ^{32}P-labelled 336-bp PCR product and the full length 1912-d05 I.M.A.G.E. DNA clone, respectively.

After exposure of the hybridized filters onto photographic film, coincident plaques were selected by alignment of the autoradiographs. For example, the exposed films for lifts 1.1 and 1.2 were aligned and only coincident positive plaques were selected from plate 1.

2.5.4.2. Secondary screening: After the primary phage was titred (from 10^{-2} to 10^{-6}), lifts from selected plates using Hybond N+ filters (Roche) were hybridized with radio-labelled 336-bp PCR product to screen for positive plaques.

2.5.5. Characterization of clones isolated from the library

Cre-expressing BNN132 cells were transformed with the isolated phage (from the secondary library screening) which resulted in automatic subcloning and isolation of the genomic insert contained within the pBluescript vector (Fig. A1c, Appendix 1, p. 168). Therefore, the successfully excised clones grew as normal bacterial colonies without the need for phage growth protocols. DH5α competent cells were also transformed using the resulting plasmid DNA.

Restriction digests of the isolated genomic DNA were prepared using a range of enzymes, loaded onto midi agarose gels and then Southern blotted. The filters were hybridized with radio-labelled 336-bp PCR product to confirm I-1 sequence homology of the clones. The gels and resulting autoradiographs were analysed by the Gel-Pro™ Analyzer for the Macintosh, Version 2.0 (Media Cybernetics, USA) computer program to give accurate estimates of the DNA fragment sizes.

2.5.6. Orientation of the inhibitor-1 gene and the restriction maps of the clones isolated after the library screening

The orientation of the I-1 gene was determined by the preparation of duplicate gels followed by hybridization using radio-labelled the 5' and 3' subclones from clone 1912-d05.

The restriction maps of clone types were determined using the restriction digest data and analysis of the gels by the Gel-Pro™ computer program.

2.5.7. Subcloning of the inhibitor-1 gene

The isolated clones from the library screening were digested with restriction enzymes, loaded onto agarose gels and Southern blotted. The resulting positive bands (i.e. after hybridization with the 1912-d05 full-length clone or the 336-bp PCR product) were subcloned into pBluescriptIIKS– vector (Fig. A1d, Appendix 1, p. 169). Using the restriction map from each clone as a guide, non-hybridizing fragments were also subcloned.

2.5.8. The sequence of the mouse inhibitor-1 gene

The subcloned genomic restriction fragments prepared in section 2.5.7 were commercially sequenced. Each subclone was sequenced using universal and reverse primers, but the sequence derived from the 3'–5' direction was rewritten to give the reverse-complement sequence. The commercial companies DNASHEF (Edinburgh, UK), MWG-Biotech (Germany) and Molecular Biology Services (King's College, London, UK) were used.

The mouse genomic sequences were analysed using the bioinformatic program NIX at the UK HGMP Resource Centre, Hinxton, Cambridge, UK (http://www.hgmp.mrc.ac.uk/).

2.5.9. Chromosomal localization of the inhibitor-1 gene

Chromosomal localization was carried out by Shelagh Boyle, MRC, Edinburgh, UK. The protocol for FISH was used as described by Fantes *et al.* (1992).

2.6. Specialized materials used in Chapter 4

3-aminopropyltriethoxysilane (TESPA) was used to coat microscope slides to prevent loss of tissue after sectioning. TESPA was handled according to the safety procedures provided from the supplier (Sigma).

PBSTX, i.e. PBS with 1% Triton X-100 (Sigma), was used as a wash and diluent for immunocytochemistry. A stock 1% solution of the preservative sodium azide was prepared and used at a final concentration of 0·01% for blocker solutions to prevent growth of bacteria.

2.7. Specialized methods used in Chapter 4

2.7.1. Preparation of TESPA slides

The microscope glass slides (BDH) were loaded into racks, and cleaned in the dishwasher, air dried, washed in absolute alcohol with 1% hydrochloric acid (or glacial acetic acid) and then air dried again.

The slides were prepared as follows:

(1) placed in 2% TESPA in acetone for 5 s;

(2) transferred to acetone for 2 × 15 s; and

(3) washed in water for 15 s then air dried.

2.7.2. Wax embedding of tissue

Tissue was fixed in 4% PFA for 1–3 h (depending upon size) at room temperature on an orbital shaker. For larger tissue (e.g. whole-mount embryos), samples were stored overnight at 4 °C in fixative to allow for adequate penetration. The tissue was washed in PBS, placed into 30% ethanol for one hour and then into 70% ethanol overnight. The sample was passed through 85%, 95% then 2 × 100% ethanol (15 min per solution). The tissue was transferred to a glass Bijoux and placed in 50% absolute alcohol:50% xylene for ~15 min. This was replaced with xylene until the tissue 'cleared'. The tissue was placed into three separate liquid wax baths (if the

sample was large, a vacuum was applied) for 10 min per bath. The tissue was then embedded in wax and stored at 4 °C for at least 24 h. The wax-embedded tissue was then transferred onto a 'chuck' or wooden block for sectioning.

2.7.3. Sectioning of tissue onto TESPA slides

Seven micrometre sections of wax embedded tissue were prepared using a microtome and placed onto the TESPA slides. The racks of slides were wrapped in foil and placed in a 37 °C oven to dry the sections. The slides were then stored at room temperature until required.

2.7.4. Inhibitor-1 expression in whole-mount tissue

The mouse embryos were supplied by John Verth, Jean Flockhart and Margaret Keighren (Department of Biomedical Sciences, University of Edinburgh). Therefore, whole-mount tissue from a range of mouse strains was analysed. The tissue was fixed in 4% paraformaldehyde (PFA) in PBS for several hours or overnight, depending upon size. Samples were then stored at 4 °C in PBSTX or PBS, and then processed as follows:

(1) The tissue was incubated in the primary anti-I-1 antibody at room temperature using an orbital shaker to ensure exposure of all of the tissue to antibody. The anti-I-1 antibodies from Serotec Ltd (Kidlington, UK) and Santa Cruz Biotechnology, Inc. (Santa Cruz, CA, USA) were diluted 1:200 and 1:50 in PBSTX, respectively.

(2) The tissue was washed in PBSTX over several hours. The PBSTX was replaced several times.

(3) The tissue was blocked in 5% normal serum or 1% BSA in PBSTX/azide. The tissue was incubated for several hours at room temperature on the orbital shaker.

(4) The tissue was placed into the FITC-labelled secondary antibody (Sigma), diluted 1:100 in PBSTX. The diluted antibody was vortexed then spun at full speed in a microcentrifuge for 5 min to remove conjugates before applying to the tissue. The samples were incubated overnight in secondary antibody on the orbital shaker.

(5) The tissue was washed in PBSTX over several hours.

(6) The tissue was washed in 2 × SSC then incubated in 10 μg mL^{-1} RNase (from a stock of 10 mg mL^{-1}) for approximately 1 h at 37 °C. During the final 20 min incubation, propidium iodide was added to a final concentration of 0·75 μg μL^{-1}. For larger tissues, the samples were incubated in RNase and propidium iodide overnight.

(7) The tissue was washed in PBS, mounted in Vectashield and analysed using the Leica TCS NT confocal microscope.

Methanol-fixed kidneys were also processed as described but the tissue was stored at –20 °C in fixative and no Triton X-100 was added to any of the solutions. Gluteraldehyde (Sigma) was added to the PFA for some of the PFA-fixed kidneys.

2.7.5. Pax-2 expression in whole-mount tissue

The tissue was fixed as described in section 2.7.4, above. However, the tissue was incubated overnight at 37 °C in anti-Pax-2 antibody (made in rabbit, diluted 1:200 in PBSTX; a kind gift from Dr Dressler). The tissue was processed as described in section 2.7.4, steps 3–5. The tissue was then incubated in secondary antibody (FITC, anti-rabbit, Sigma) diluted 1:100 and processed as described above.

2.7.6. Laminin expression in whole-mount tissue

The tissue was processed as described for the I-1 antibody (see section 2.7.4, above) but an anti-laminin primary antibody (Sigma) and an anti-rabbit FITC secondary antibody (Sigma) were used.

2.7.7. Pan cytokeratin expression in whole-mount kidney

Kidneys were fixed and stored in methanol at –20 °C until required. Kidneys were processed as described for the I-1 antibody (see section 2.7.4, above) but used anti-

pan cytokeratin (Sigma) and anti-rabbit FITC secondary (Sigma) antibodies. No Triton X-100 was added to any of the solutions.

2.7.8. Inhibitor-1 expression in kidney sections

(1) All incubations were carried out at room temperature. The sections were de-waxed for 10 min each in xylene; xylene and absolute alcohol (50/50); and then ethanol dilutions of 95%, 85%, 70%, 50% and 30%. The slides were then incubated in PBS for 5 min. The sections were blocked in serum or 1% BSA (diluted in PBSTX) for ~1 h.

(2) The blocker was tapped off the slide and any remaining solution was carefully dried around the section.

(3) The I-1 antibody (Serotec) was diluted 1:500 in PBSTX and applied under a coverslip and incubated overnight in a moisture chamber.

(4) The slides were washed for 2 × 15 min in PBSTX.

(5) The sections were blocked for ~1 h.

(6) The blocker was tapped off the slide and carefully dried around the section.

(7) The anti-sheep FITC secondary antibody was diluted 1:50 in PBSTX and was applied under a coverslip, and incubated for over 1 h at room temperature.

(8) The slides were washed for 2 × 15 min in PBSTX.

(9) The FITC-labelled sections were mounted in Vectashield.

2.8. Specialized materials used in Chapter 5

The medium used for tissue culture contained L-glutamine and was ordered from Sigma. The foetal calf serum (FCS), sterile disposable 3-mL plastic pastettes, Iwaki tissue culture plasticware and the individual 3-cm^2 sterile plastic dishes were all ordered from Labtech International (Ringmer, UK). The cell lines were cultured according to ECCAC method sheets. The FuGENE^{TM} 6 transfection reagent (Roche) was used according to manufacturer's instructions.

2.9. Specialized methods used in Chapter 5

2.9.1. Modification of genomic type 2 clone and pBluescriptII

The insertion of oligonucleotides (to introduce unique restriction sites) into clone type 2 was carried out according to the protocol described by Doyle (1996). The oligonucleotides were modified by the addition of a 5' phosphate group and were synthesized commercially (MWG-Biotech). The complementary oligonucleotides were heated together in a boiling waterbath for 5 min, cooled to room temperature to anneal and were then ligated to the genomic DNA. If the oligonucleotides were to be inserted into a site that was not unique, then the DNA was partially digested. The resulting modified DNA was used to transform competent DH5α cells.

The vector pBluescriptIIKS–/+ was digested with *Sal*I and then the site was removed using the enzyme *Pfu* polymerase (Stratagene). This enzyme fills in the 5' overhangs in the presence of excess dNTPs. After treatment, the pBluescriptIIKS vector was re-ligated and used to transform competent DH5α cells. The plasmid vector was then digested to ensure that the restriction site had been deleted. Genomic DNA was then inserted into this modified pBluescriptII vector.

2.9.2. The pGT1·8IRESβgeo construct

The pGT1·8IRESβgeo vector (Fig. A1e, Appendix 1, p. 170) was digested, isolated by agarose gel electrophoresis and GENECLEANed. The genomic DNA (inserted into the *Sal*I knock-out pBluescriptII vector) was digested then treated with CIAP. Ligation reactions were set up to insert the GT1·8IRESβgeo DNA into the genomic plasmid.

2.9.3. The pEGFP-1 construct

The modified type 2 clone (from section 2.9.1, above) was digested then separated by agarose gel electrophoresis to isolate the desired fragment. The DNA was ligated into the multiple cloning site (MCS) of the pEGFP-1 vector (Clontech, Palo Alto

USA; GenBank accession number U55761; Fig. A1f, Appendix 1, page 171). A shorter genomic restriction fragment was also ligated into the MCS of pEGFP-1

2.9.4. Preparation of DNA for use in transfection experiments

Maxipreps of the pEGFP-1 vector only, pEGFP-N1 (Clontech; accession number U55762; Fig. A1g, Appendix 1, page 172) and the genomic constructs were prepared using the Promega Wizard® *Plus* MaxiPrep DNA Purification System.

2.9.5. Cell culture

HEK293 cells (a kind gift from Dr Rory Duncan, Department of Biomedical Sciences, University of Edinburgh) were grown in Eagle's minimum essential medium with 10% foetal calf serum. 4/4 RM-4 cells were grown in Ham's F12 medium with 10% foetal calf serum, according to the information sheet supplied by ECACC. All cells were grown at 37 °C in a humidified CO_2 incubator.

For the transfection studies, a confluent 25-cm^2 tissue culture flask was trypsinized and the cells were spun for 5 min at 1000 rpm, then resuspended in 10 mL of the appropriate medium. A sterile, autoclaved 22×22 mm glass coverslip was placed at the bottom of a 30-mm Petri dish before the addition of 1 mL of the cell suspension. A further 1 mL of medium was added to each plate and this gave an approximate seeding of 1×10^5 cells per plate.

2.9.6. Inhibitor-1 protein expression of the cell lines

The cell lines were cultured in individual 30-mm plates, as described in section 2.9.5. The cells were grown overnight or until confluent, fixed in 4% PFA for 20 min at room temperature and incubated with the anti-I-1 antibody (Santa Cruz Biotechnology, Inc.; diluted 1:50 with PBSTX) for several hours. After washing in PBSTX, the cells were blocked with serum then incubated in anti-goat FITC secondary antibody (diluted 1:200) for several hours. The cells were washed then the

coverslip was mounted in Vectashield (containing propidium iodide at a concentration of 0·75 μg μL^{-1}) onto a microscope slide, sealed with nail varnish and then viewed using the confocal microscope. Control cells were not incubated with the I-1 antibody.

2.9.7. Transfection of the cell lines

The transfection protocol outlined in the Roche handbook was followed. Fixed cells were washed in PBS and treated with 4% PFA for 20 min at room temperature, whereas unfixed cells were washed in PBS only before mounting. The coverslips were mounted as described in section 2.9.6.

The number of transfected cells was determined by counting the positive cells per 500×500-μm area using the confocal microscope computer screen. The average number of positive cells was taken from six random 500 μm^2 areas across the slide. Each transfection was repeated at least twice to ensure reproducibility of results.

Chapter 3

Isolation and genomic characterization of the mouse inhibitor-1 gene

3.1. Introduction

3.1.1. Abstract

A range of mouse cDNA clones which contained sequence homology to rat I-1 mRNA were selected from a database after a keyword search. One of these clones (1912-d05) was shortened to give a 5' PCR product and a 3' restriction fragment. A mouse genomic library was screened using the 5' PCR product of 1912-d05 and two types of clones were isolated from the library that contained homology to rat I-1 mRNA. Sequencing and analysis of these two clone types showed that the protein coding region of the mouse I-1 gene spans ~7 kb of genomic DNA and is encoded by seven exons. The predicted mouse mRNA and protein sequences are 95% homologous to the published rat sequences, and the I-1 gene is located on chromosome 15, band F.

3.1.2. Aim of the chapter

The aim of this chapter was to isolate and characterize the mouse protein phosphatase I-1 gene. Analysis of the known I-1 sequences of rabbit, rat and human shows strong homology, and therefore, it was thought likely that the mouse I-1 sequence was similar to those published. The plan of work for this research chapter was to screen a mouse genomic library to isolate clones which shared sequence homology to rat I-1 mRNA. The recombinants isolated after the screening could then be characterized by restriction digestion, and these clones would then be used to produce transgenic mice (Chapter 5). After the recombinants had been subcloned and sequenced to confirm I-1 homology to the rat I-1 mRNA, then the predicted I-1 mRNA and protein sequences could be derived. The chromosomal localization of the I-1 gene could also be determined by fluorescent *in situ* hybridization (FISH) and compared to other genes found in this chromosome region.

3.1.3. Strategy for isolating the mouse inhibitor-1 gene

The UK Medical Research Council (MRC) formed the Human Genome Mapping Project Resource Centre (HG-MPRC) in 1990. The centre provides facilities for genome mapping and gene isolation for the research community, as well as access to various databases and biological materials (e.g. human YAC and mouse cosmid libraries). Access to the facility was obtained through the Internet (http://www.hgmp.mrc.ac.uk/), online help and advice was available if required. One of the databases was screened using a keyword search and a list of mouse cDNA clones homologous to the rat I-1 mRNA sequence was obtained, allowing rapid and reliable isolation of clones relevant to the project. One of these mouse cDNA clones was selected for the library screening, but was shortened to give 5' and 3' fragments.

A cDNA library is formed from messenger RNA (mRNA) isolated from tissue at a particular developmental age and only contains the transcripts of expressed genes, unlike a genomic library, which represents the entire genome. The commercial library used in this project was a λPS vector that contained 16–20-kb inserts of mouse genomic DNA. There were two loxP sites flanking the insert that was targetted for recombination, and in the presence of the enzyme Cre recombinase, there was excision of the insert from the phage backbone (Fig. A1b, Appendix 1, p. 167). The presence of pBluescript in the insert allowed easier handling of the final isolated clones. The Cre-loxP system allows tissue- and time-specific DNA recombination and has been used *in vivo* (for review, see Stricklett *et al.* 1999). Screening of the library using the 5' fragment of a selected cDNA clone reduced the potential for the selection of false positive clones from the library since the homopolymeric 3' sequence was eliminated. The orientation of the I-1 gene was also determined by hybridization of identical Southern blots using the ^{32}P-radio-labelled 5' and 3' DNA fragments.

Restriction digestion allowed the restriction map for each isolated clone to be determined. Southern blotting also confirmed the presence of I-1 gene homology of the clones isolated after library screening and allowed the DNA to be subcloned then sequenced. The sequences of these positive clones were determined,

followed by comparative analysis using the computer program NIX, which is a tool to aid the identification of nucleic acid sequences (http://menu.hgmp.mrc.ac.uk/menu-bin/Nix/Nix.pl). NIX encompasses a variety of different nucleic acid analysis computer programs, but has the advantage that the sequence of interest is submitted only once. As a final characterization of the gene, the chromosomal localization of I-1 was visualized by FISH.

3.2. Results

3.2.1. I.M.A.G.E. clone analysis

The first step towards cloning of the mouse I-1 gene was to obtain mouse cDNA clones with homology to the rat I-1 mRNA sequence, one of which would then be used to screen the mouse genomic library. Integrated Molecular Analysis of Genomes and their Expression (I.M.A.G.E.; Lennon *et al.* 1996) had a variety of high quality, partially sequenced clones, and the sequencing data had been deposited in the public databases (e.g. EMBL and GenBank). Only the first few hundred base pairs of the clones had been sequenced by I.M.A.G.E., and therefore, there was only partial homology to the deduced rat I-1 mRNA. Ten candidate Expressed Sequence Tags (ESTs) were identified after a keyword search of the dbEST database using the Entrez computer program. This facility was available from the MRC HGMP-RC website at http://www.hgmp.mrc.ac.uk/. The partial sequences were submitted to the sequence database FASTA (FastA@ebi.ac.uk) to determine the extent of homology to rat I-1 mRNA (Table 3.1) and the 10 selected clones were ordered through HGMP-RC.

Table 3.1. The I.M.A.G.E. clones selected from the dbEST database. The dbEST identification (ID) number, clone name and I.M.A.G.E. ID number and percentage homology to rat I-1 mRNA are listed.

DbEST ID	Clone Name	I.M.A.G.E. ID number	Percentage homology to rat I-1 mRNA (FASTA)
1562249	3174-c15	1260998	94·9
1156220	2313-c12	0930371	94·3
745267	1244-p24	0520199	95·0
1481105	3141-o09	1248608	92·9
517532	706-j19	0313458	95·0
926702	1912-d05	0776404	94·1
673408	804-h12	0351035	94·7
644539	1099-j07	0464358	60·2
674484	940-i01	0403272	95·0
761351	1396-a05	0578188	93·9

The I.M.A.G.E. clones were received and the plasmid DNA isolated according to section 2.3.6 in Chapter 2. Preliminary digestion of the clones determined insert size (Table 3.2).

Table 3.2. The I.M.A.G.E. clones obtained from HGMP-RC and the insert size determined by restriction digest and agarose gel electrophoresis: NA–the insert size was not stated by I.M.A.G.E.

Clone name	Vector	Enzymes used to isolate insert from vector	Insert size stated by I.M.A.G.E.	Insert size predicted by gel (kb)
3174-c15	pBluescript	EcoR1, Xho	0·6	1
2313-c12	pBluescript	EcoR1, Xho	1·5	1·5
1244-p24	pBluescript	EcoR1, Xho	1	1
3141-o09	pT7T3D-Pac	EcoR1, Not1	NA	1
706-j19	pT7T3D-Pac	EcoR1, Not1	NA	0·8
1912-d05	pT7T3D-Pac	EcoR1, Not1	0·7	0·8
804-h12	pT7T3D-Pac	EcoR1, Not1	NA	1
1099-j07	pT7T3D-Pac	EcoR1, Not1	NA	1
940-i01	pT7T3D-Pac	EcoR1, Not1	NA	0·8
1396-a05	pT7T3D-Pac	EcoR1, Not1	NA	1·2

Digestion of the I.M.A.G.E. DNA using the appropriate enzymes listed in Table 3.2 showed the presence of insert. However, clones 3174-c15, 2313-c12 and 1244-p24 gave unpredicted fragments when digested with *Eco*RI and *Xho*I, and thus, further analysis of these was abandoned. Some of the I.M.A.G.E. clone inserts had a higher molecular weight than stated in the data sheet from I.M.A.G.E.

Clones 1912-d05 and 3141-o09 were randomly selected for further study and were sequenced to confirm identity (Figs A2–A4, Appendix 1, pp. 173–178). For further analysis, the shorter clone 1912-d05 was selected since this clone would have a reduced 3' homopolymeric sequence in comparison to clone 3141-o09. Clone 1912-d05 was derived from a mouse BALB/c kidney library (developmental age 6 weeks) constructed by Dr Robert Barstead as part of the Marra M/Mouse EST project (unpublished results 1996).

3.2.2. Subcloning of the 5' PCR product and 3' restriction fragment of 1912-d05 into vectors

A 5' PCR product and a 3' restriction fragment were generated from clone 1912-d05. The main reason for this was that the full-length and 5' product of clone 1912-d05 would be used for screening the genomic library and only coincident plaques would be selected: the presence of the 3' homopolymeric region may give rise to false positive clone selection. Also, orientation of the I-1 gene could be determined using the 5' and 3' fragments hybridized to duplicate gels (see section 3.2.6, below).

The Webcutter computer program (http://www.medkem.gu.se) confirmed that there was a lack of suitable restriction sites within the 1912-d05 sequence. Therefore, PCR was used to generate a 5' product since the size of the fragment could be accurately predicted. The primers were designed to allow a substantial 5' fragment to be produced and the final 336-bp PCR product sequence was homologous to the sequences of rat I-1 mRNA, and the I.M.A.G.E. 1912-d05 and 3141-o09 sequences (Fig. 3.1).

Figure 3.1. The first 400-nucleotide I.M.A.G.E. sequence of clone 1912-d05. The positions of the forward and reverse primers are indicated by the highlighted green and pink nucleotides, respectively. Thus, a 336-bp product was generated.

```
  1  GTTCAGAGGC CAAGAATTCG GATCCATGGC CGCTGCCTGA CCGGGAGCCA
 51  TGGAGCCCGA CAACAGCCCA CGGAAGATCC AGTTTACGGT CCCGCTGCTG
101  GAGCCTCACC TGGACCCGGA GGCGGCGGAG CAGATTCGGA GGCGCCGCCC
151  CACCCCTGCC ACACTTGTGC TGACCAGTGA TCAGTCCTCC CCAGAAATAG
201  ATGAAGACCG GATCCCCAAC TCACTTCTCA AGTCCACCTT GTCAATGTCT
251  CCACGGCAAC GGAAGAAGAT GACAAGGACC ACACCCACCA TGAAAGAGCT
301  CCAGACGATG GTTGAACATC ACCTAGGGCA ACAGAAGCAA GGGGAAAGAA
351  CCTGAGGGAG CCACTGAGAG CACAGGGAAC CAGGAGTCCT GCCCACCTGG
```

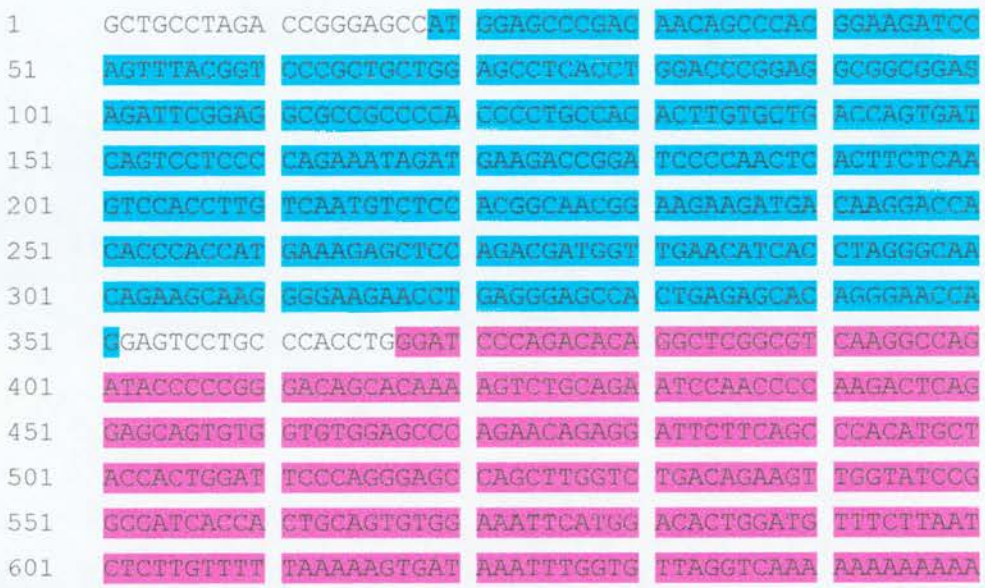
Forward primer: Nucleotides 48–66 5'–CCATGGAGCCCGACAACAG–3'

Reverse primer: Nucleotides 365–383 5'–CTGGTTCCCTGTGCTCTCA–3'

The primers were analysed using the computer program from Williamstone Enterprises at http://www.williamstone.com ('Primers! for the web'). This program predicted any potential self-annealing. The PCR product was generated using Promega 'Ready-To-Go' PCR beads, and the product was excised from an agarose gel, GENECLEANed and inserted into pGEM®-T Easy vector (Fig. A1a, Appendix 1, p. 166).

A triple digest of 1912-d05 DNA with *Not*I, *Bam*HI and *Eco*RI isolated a *Bam*HI/*Not*I 3' fragment (~280 bp) that was excised from a low melting point gel then ligated into pBluescriptIIKS–. Figure 3.2 shows the sizes of the 5' PCR product and the 3' restriction fragment in relation to each other. The 3' fragment was ~280 bp in length, but the precise size was difficult to determine because of degeneration of sequencing towards the 3' end.

Figure 3.2. The sequence of 1912-d05 (from MWG-Biotech) showing the relative sizes of the 5' PCR product and the 3' fragment highlighted in blue and pink, respectively.



Gel electrophoresis and sequencing confirmed that the 5' 336 bp-PCR product and the 3' restriction fragment derived from clone 1912-d05 were successfully ligated into the vectors (Figs A1a & d, Appendix 1, pp. 166 & 169, respectively).

3.2.3. Library preparation

The genomic λPS library (male 129 SV D3, MoBiTec) was screened in order to isolate genomic clones which contained homology to rat I-1 mRNA sequence. Initially, the titre of the library was determined by diluting the library stock from 10^{-2} to 10^{-6}. Since each colony or plaque was derived from a single infected bacterial cell, the average number of plaque forming units (PFUs) per microlitre of library stock was calculated (Table 3.3, below). As a control, water replaced the λ DNA.

Table 3.3. The number of colonies per dilution of the λPS library: (PFUs) plaque forming units.

Dilution of library	Number of PFU/plate	Average number of PFU/μl stock of library
Control	0	0
10^{-6}	0	0
10^{-5}	12	10^{6}
10^{-4}	116	10^{6}
10^{-3}	980	10^{6}
10^{-2}	Confluent	–

In accordance with the protocol from MoBiTec, 2×10^{6} phage allowed complete representation of the mouse genome. Therefore, 2 μL of the stock library was plated and the resulting filters were processed as described in Chapter 2.

3.2.4. Screening of the library

The first and the second lifts of the library were hybridized with the ^{32}P-labelled 336-bp PCR product and the full-length 1912-d05 I.M.A.G.E. DNA clone, respectively. After primary screening, a total of 36 plaques were picked (Table 3.4). Identification of the plugs refers to the plate number followed by the plug picked from the plate e.g. 1A referred to plate number 1, plug A. Only coincident plaques were selected to reduce the potential for negative clone selection.

Table 3.4. The plugs selected after the primary screening of the λPS library; for example, 1A refers to plate 1, plug B; and 7B refers to plate 7, plug B.

1A; 1B; 1C; 1D; 1E; 1F; 1G; 1H
2A; 2B
3A
4A; 4B; 4C; 4D; 4E; 4F; 4G
5A; 5B
6A
7A; 7B; 7C; 7D
8A; 8B; 8C; 8D; 8E
9A; 9B; 9C; 9D
10A; 10B

After the primary phage plug was titred (using the range 10^{-2} to 10^{-6}), lifts were taken from selected plates using Hybond N+ filters (Roche). These filters were hybridized with the radio-labelled 336-bp PCR product to screen for positive plaques. After this secondary screening, 60 positive clones were isolated (see Table 3.5 for the full list). A tertiary screening was not necessary since the positive plugs selected were well isolated, which reduced the chances of contamination from neighbouring plaques.

Table 3.5. The plugs selected after the secondary screening of the λPS library; for example, 1A1 refers to plug 1 selected from plate 1A; and 8D6 is plug number 6 selected from plate 8D.

1A1; 1A2; 1A3; 1A4; 1A5
1C1; 1C2; 1C3; 1C4; 1C5
1D1; 1D2; 1D3; 1D4; 1D5; 1D6; 1D7; 1D8; 1D9; 1D10; 1D11
3A1; 3A2; 3A3; 3A4; 3A5; 3A6; 3A7; 3A8; 3A9
4A1; 4A2; 4A3; 4A4
4C1; 4C2; 4C3; 4C4; 4C5; 4C6
8D1; 8D2; 8D3; 8D4; 8D5; 8D6; 8D7; 8D8; 8D9; 8D10
9A1; 9A2
9D1; 9D2; 9D3; 9D4; 9D5; 9D6; 9D7; 9D8

3.2.5. Characterization of clone types isolated from the library

The isolated clones had DNA inserts (from 16 to 20-kb) packaged in a phage vector. The λPS vector contained an automatic plasmid subcloning facility that allowed isolation of large fragment sizes of genomic DNA with the benefits of plasmid handling. Within the phage, there were two loxP sites flanking the pBluescript sequence and insert. Initially, the phage were transferred into *E. coli* C600 plating cells (these cells do not express Cre recombinase). When phage were introduced into the Cre-recombinase expressing cells *E. coli* BNN132, automatic subcloning resulted in a loss of phage material to reveal pBluescript plasmid containing the genomic insert (Fig. A1b, Appendix 1, p. 167). This allowed ease of handling and manipulation of the resulting genomic DNA clones, and was more convenient than working with standard phage libraries.

The following recombinants isolated from the library were selected and BNN132-competent cells were transformed to isolate the insert DNA contained within pBluescript (Fig. A1c, Appendix 1, p. 168):

1A1A 1A1B 1C1A 1C1B 1D1A 1D1B 3A1A 3A1B 4A1A 4A1B
4C1A 4C1B 8D1A 8D1B 9A1A 9A1B 9D1A 9D1B

The yield of DNA after transformation of BNN132 cells was very low. However, DH5α competent cells were transformed with the plasmid DNA, and the resulting DNA was more concentrated and there was no loss of genomic DNA, as confirmed by restriction digest and agarose gel electrophoresis.

The plasmid DNA isolated from the clones was double digested with *Kpn*I and *Sac*I to release the insert DNA from pBluescript vector, and four different types of clones were determined by the restriction digest pattern, termed types 1–4. In order to further characterize the clones, PCR analysis was carried out using Promega Ready-To-Go beads (following the instruction manual to optimize the reaction conditions) using the forward and reverse primers designed in section 3.2.2. However, analysis of the PCR products by agarose gel electrophoresis showed the presence of only three different types of clones (Table 3.6).

Table 3.6. The different types of clones (1–4) isolated after screening of the λPS library, as determined by double restriction digest with *Kpn*I and *Sac*I, and PCR.

Clone name	KpnI/SacI digest	PCR pattern
1A1A	1	1
1A1B	1	1
1C1A	1	1
1C1B	1	1
1D1A	1	1
1D1B	1	1
3A1A	1	1
3A1B	1	1
4A1A	2	2
4A1B	2	2
4C1A	3	3
4C1B	1	1
8D1A	1	1
8D1B	1	1
9A1A	4	3
9A1B	4	3
9D1A	1	1
9D1B	1	1

The pooled restriction digestion and PCR results suggested that there were four different types of clone isolated from the library. Clones 4C1A, 9A1A and 9A1B all had the same PCR pattern (type 3), but when double digested with *Kpn*I and *Sac*I, clone 4C1A generated a different restriction pattern to 9A1A and 9A1B.

DNA from clones 1C1A, 4A1A, 4C1A and 9A1A (selected to represent clone types 1–4, respectively) was digested with a range of double digests, and the midi gels were blotted and hybridized with radio-labelled 336-bp PCR product. This further characterization showed that clone types 4C1A and 9A1A (types 3 and 4, respectively) did not have any sequence homology to I-1 since the resulting autoradiographs were blank. However, DNA from clones 1C1A and 4A1A (types 1 and 2, respectively) had I-1 homology (Figs 3.3 & 3.4).

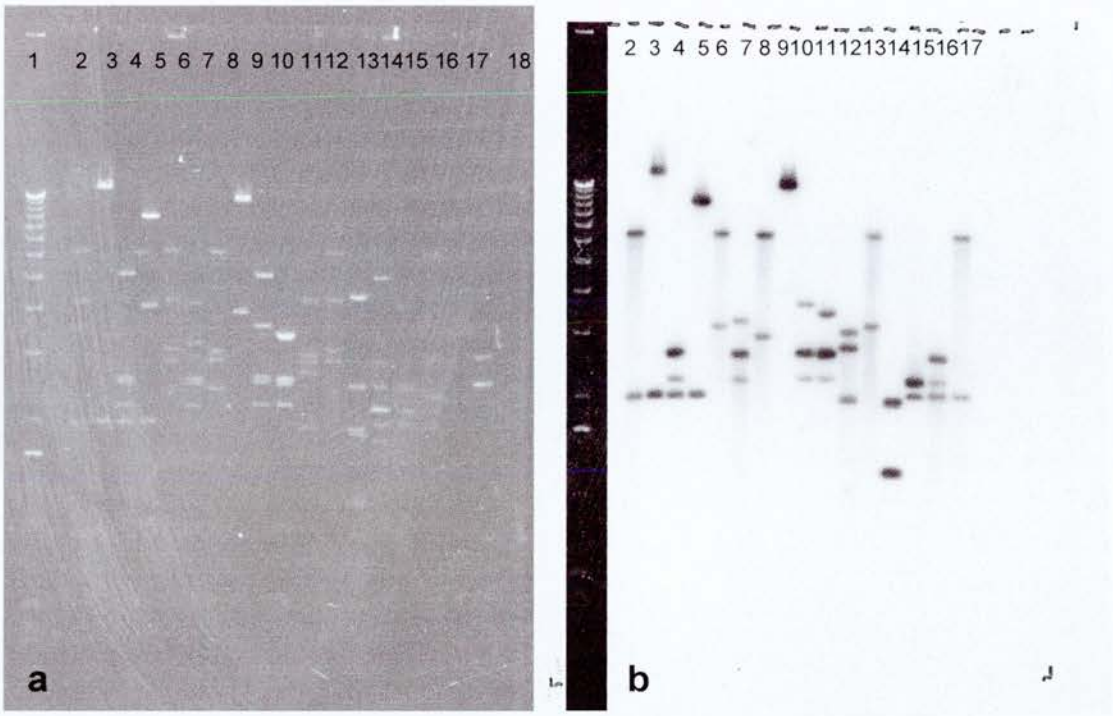


Figure 3.3. Analysis of type 1 clone DNA: (a) 1 × TBE, 0·7% agarose gel of type 1 clone DNA (1C1A) digested with a range of enzymes; and then (b) Southern blotted and hybridized with the 5' 336-bp PCR product. The gel and autoradiograph was analysed by the Gel-Pro computer program, and the 1-kb ladder from the agarose gel was superimposed onto the autoradiograph.

Lanes:

Lane 1: 1-kb ladder, 1 μg
Lane 2: *Hind*III*Kpn*I
Lane 3: *Hind*III*Not*I
Lane 4: *Hind*III*Pvu*II
Lane 5: *Hind*III*Xba*I
Lane 6: *Kpn*I*Not*I
Lane 7: *Kpn*I*Pvu*II
Lane 8: *Kpn*I*Xba*I
Lane 9: *Not*I*Xba*I
Lane 10: *Not*I*Pvu*II
Lane 11: *Pvu*II*Xba*I
Lane 12: *Kpn*I*Sac*I
Lane 13: *Kpn*I*Not*I
Lane 14: *Bam*HI*Pvu*II
Lane 15: *Pvu*II*Eco*RI
Lane 16: *Hind*III*Kpn*I*Pvu*II
Lane 17: *Hind*III*Kpn*I*Xba*I
Lane 18: 50-bp ladder, 500 ng.

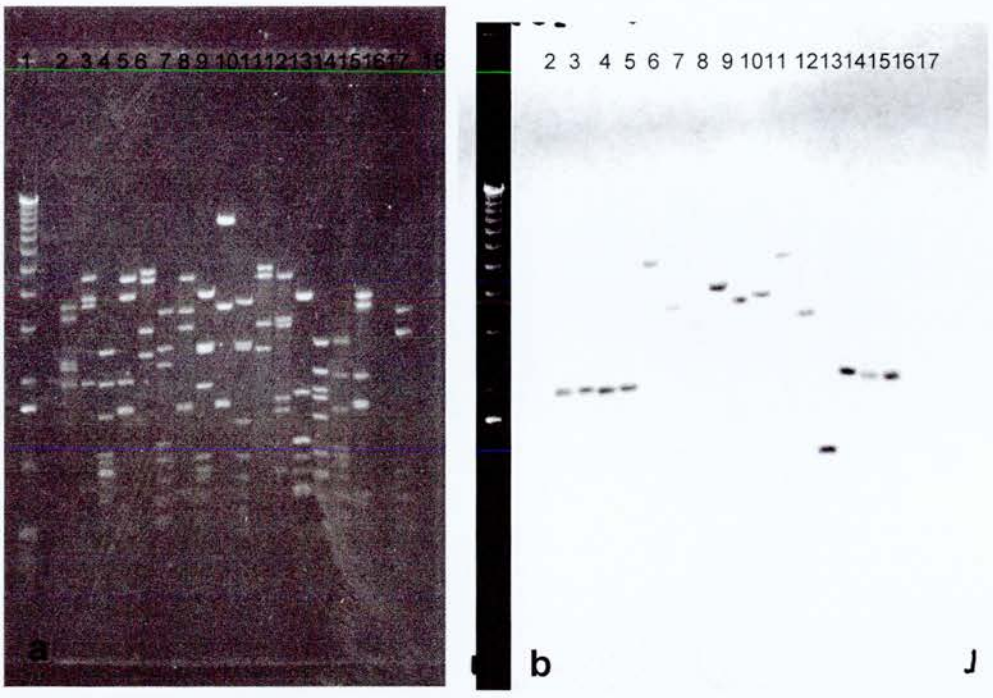


Figure 3.4. Analysis of type 2 clone DNA: (a) 1 × TBE, 0·7% agarose gel of type 2 clone DNA (4A1A) digested with a range of enzymes; and then (b) Southern blotted and hybridized with the 5' 336-bp PCR product. The agarose gel and the autoradiograph were analysed by the Gel-Pro computer program and the 1-kb ladder superimposed onto the autoradiograph. Lanes:

Lane 1: 1-kb ladder, 1 μg
Lane 2: *Hind*III*Kpn*I
Lane 3: *Hind*III*Not*I
Lane 4: *Hind*III*Pvu*II
Lane 5: *Hind*III*Xba*I
Lane 6: *Kpn*I*Not*I
Lane 7: *Kpn*I*Pvu*II
Lane 8: *Kpn*I*Xba*I
Lane 9: *Not*I*Pvu*II
Lane 10: *Not*I*Xba*I
Lane 11: *Pvu*II*Xba*I
Lane 12: *Kpn*I*Not*I
Lane 13: *Kpn*I*Sac*I
Lane 14: *Bam*HI*Pst*I
Lane 15: *Pvu*II*Eco*RI
Lane 16: *Hind*III*Kpn*I*Pvu*II
Lane 17: *Hind*III*Kpn*I*Xba*I
Lane 18: 50-bp ladder, 500 ng

Analysis of the autoradiograph of the remaining 14 clones double-digested with *Kpn*I and *Sac*I (Fig. 3.5) showed that there was no DNA present on the gel for clone 1A1B (reason unknown). The restriction pattern for 9D1B was odd, although there was I-1 homology since the autoradiograph showed positive signal (Fig. 3.5) and this result may have been as a result of partial digestion of the clone. The previous *Kpn*I/*Sac*I digestion pattern showed that both clones 1A1B and 9D1B were of type 1 clone (Table 3.6, above).

Analysis of the *Kpn*I/*Sac*I digest of type 1 clone (Figs 3.3–3.5) gave seven restriction fragments (Table 3.7). In Fig. 3.5, the 3-kb and 2·7-kb bands were not well resolved after agarose electrophoresis, but Fig. 3.3 showed that three out of the six genomic restriction fragments had I-1 sequence homology, as shown by the presence of positive bands on the Southern blot.

Table 3.7. The restriction fragments from clone 1C1A (type 1) after digestion with *Kpn*I and *Sac*I: (+ and –) whether the fragment contained I-1 sequence homology (from Southern blot analysis); and (*) pBluescript vector.

Fragment size (kb)	Positive band on Southern blot	Fragments a result of digestion with:
4·4	–	*Kpn*I
3	+	*Sac*I*Kpn*I
2·9*	–	*Sac*I*Kpn*I
2·7	+	*Sac*I
2	+	*Sac*I*Kpn*I
0·9	–	*Sac*I
0·6	–	*Kpn*I*Sac*I

This section of work showed that the type 1 clone was the dominant recombinant isolated from the library: out of 18 clones isolated from the library and analysed by restriction digest, 13 were of type 1 and two recombinants were of type 2. The insert sizes of type 1 and type 2 clone were ~14 and ~12 kb, respectively.

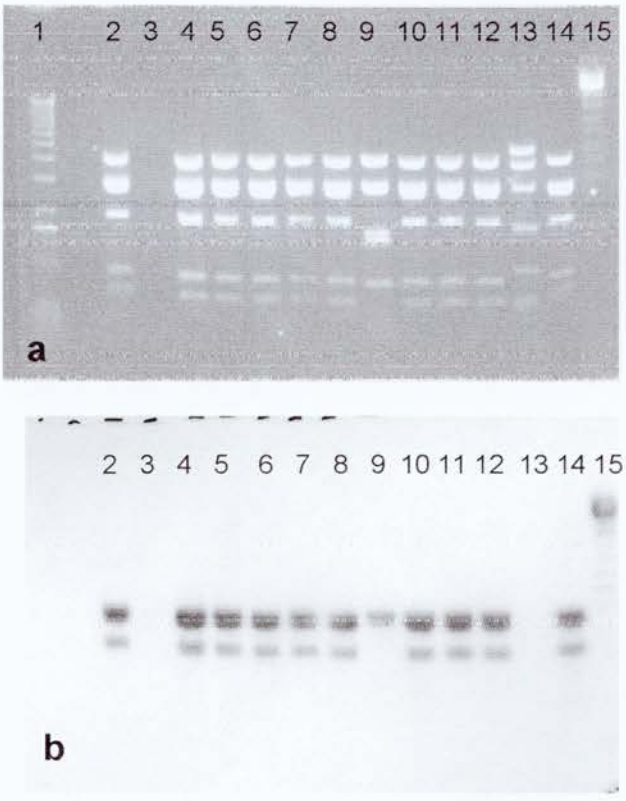


Figure 3.5. Analysis of the remaining clones isolated from the library screening: (a) 1 × TBE, 0·7% agarose gel, with the remaining clones isolated from the library screening after digestion with *Kpn*I and *Sac*I; and then (b) Southern blotted and hybridized with the 5' 336-bp PCR product. The agarose gel and autoradiograph were analysed by the Gel-Pro computer program. Lanes:

Lane 1: 1-kb ladder, 1 μg
Lane 2: 1A1A
Lane 3: 1A1B
Lane 4: 1C1B
Lane 5: 1D1A
Lane 6: 1D1B
Lane 7: 3A1A
Lane 8: 3A1B
Lane 9: 4A1B
Lane 10: 4C1B
Lane 11: 8D1A
Lane 12: 8D1B
Lane 13: 9A1B
Lane 14: 9D1A
Lane 15: 9D1B

3.2.6. Orientation of the inhibitor-1 gene and the restriction maps of the two clone types

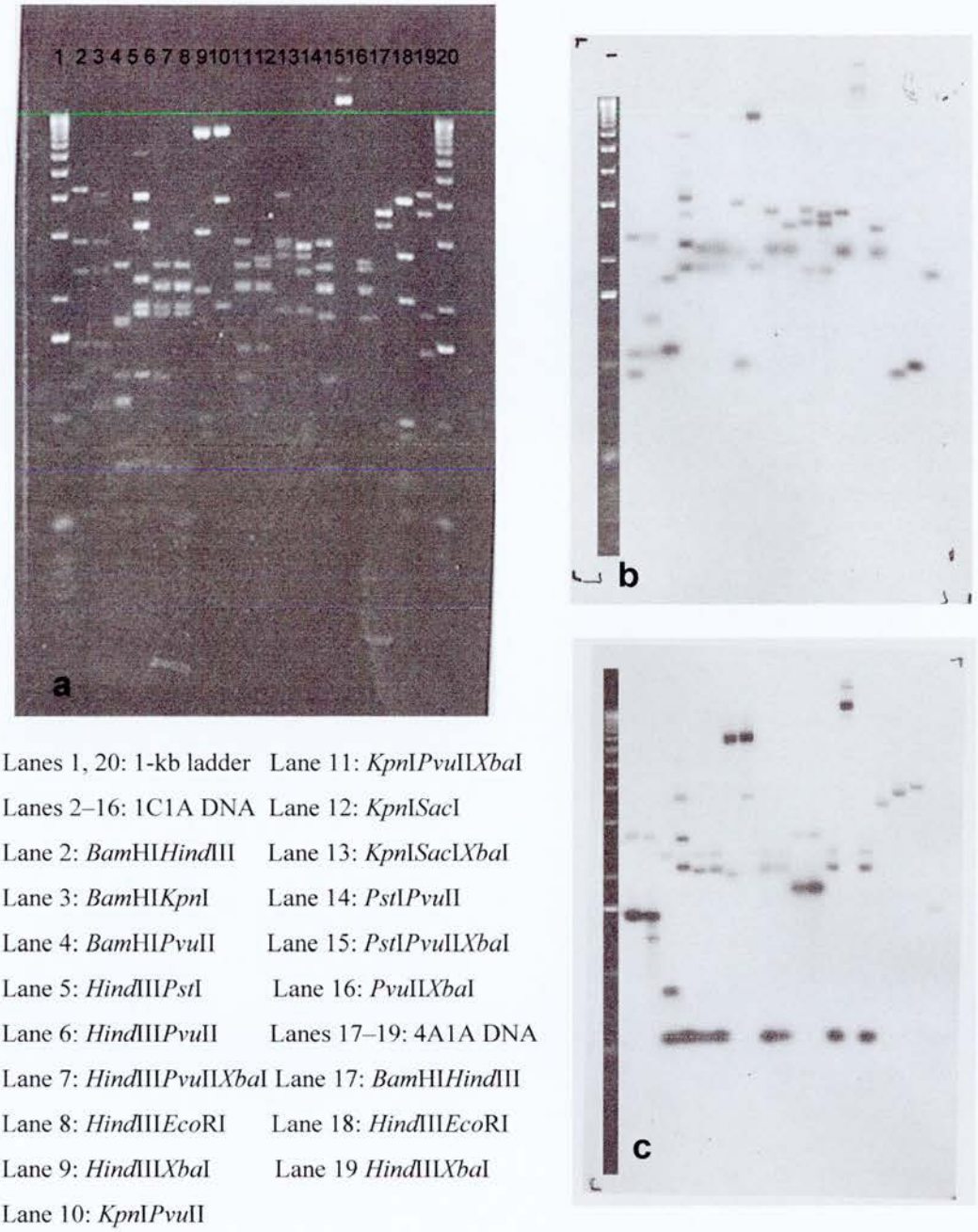
The orientation of the I-1 gene was determined after hybridization of two duplicate Southern blots with the 5' and 3' fragments of I.M.A.G.E. clone 1912-d05 (Fig. 3.6). Analysis of the *Kpn*I/*Sac*I and *Kpn*I/*Sac*I/*Xba*I digests gave the following sized bands, as shown in Table 3.8, below.

Table 3.8. The size of the *Kpn*I/*Sac*I restriction fragments (bp) from the A5 and B3 autoradiographs for type 1 (1C1A) DNA, derived from the Gel-Pro™ computer analysis system. Filters A and B were hybridized with the 336-bp PCR product and the 3' restriction fragment, respectively. The bands which are bracketed represent weak or faint bands; (–) an absent band; and bold font represents the strongest radio-labelled fragment.

***Kpn*I/*Sac*I**		***Kpn*I/*Sac*I/*Xba*I**	
A5	**B3**	**A5**	**B3**
3015	—	**2903**	—
2759	(2799)	2742	(2779)
(1982)	**1901**	(1950)	**1901**

Analysis of filter 'A' (probed with the 5' sequence) showed that the strongest hybridizing *Kpn*I/*Sac*I fragment was the 3015-bp band, followed by the 2759-bp band. The 1982-bp fragment labelled the weakest (Table 3.8 and Fig. 3.6). This suggested that the subclone at the 5' position of the gene was the 3-kb fragment, and the 3' end of the gene contained the 1·9-kb band. Therefore, the 2·7-kb band must lie between these two fragments.

Analysis of filter 'B' (probed with the 3' fragment) showed that the 1901-bp band hybridized most strongly and the 2799-bp band most weakly. The 3000-bp band did not hybridize at all (Table 3.8 and Fig. 3.6). This suggested that the fragment at the 3' end of the gene is the 1901-bp band, with the 2799-bp fragment between the other two since it weakly hybridized. Since the 3-kb fragment did not hybridize at all to the radio-labelled 3' restriction fragment, then this subclone did not contain any 3' sequence, and thus, must contain only 5' sequence. This model fitted with the results from filter A probed with the 336-bp PCR product, above.

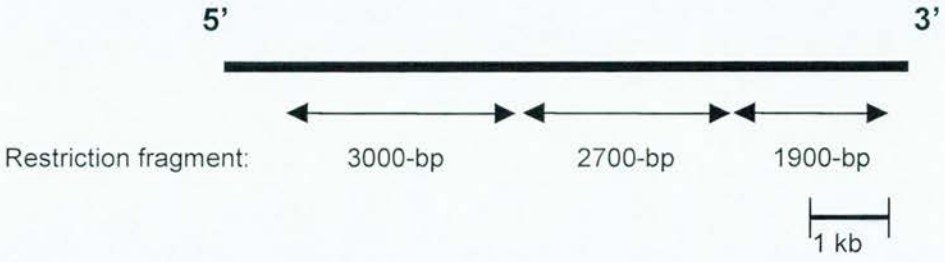


Lanes 1, 20: 1-kb ladder
Lanes 2–16: 1C1A DNA
Lane 2: *Bam*HI*Hind*III
Lane 3: *Bam*HI*Kpn*I
Lane 4: *Bam*HI*Pvu*II
Lane 5: *Hind*III*Pst*I
Lane 6: *Hind*III*Pvu*II
Lane 7: *Hind*III*Pvu*II*Xba*I
Lane 8: *Hind*III*Eco*RI
Lane 9: *Hind*III*Xba*I
Lane 10: *Kpn*I*Pvu*II
Lane 11: *Kpn*I*Pvu*II*Xba*I
Lane 12: *Kpn*I*Sac*I
Lane 13: *Kpn*I*Sac*I*Xba*I
Lane 14: *Pst*I*Pvu*II
Lane 15: *Pst*I*Pvu*II*Xba*I
Lane 16: *Pvu*II*Xba*I
Lanes 17–19: 4A1A DNA
Lane 17: *Bam*HI*Hind*III
Lane 18: *Hind*III*Eco*RI
Lane 19 *Hind*III*Xba*I

Figure 3.6. The orientation of the I-1 gene determined by restriction digest followed by agarose gel electrophoresis and Southern blot. (a) Two identical 1 × TBE, 1·2% agarose gels with DNA digests of clone types 1 and 2. The Southern blots were then hybridized with the (b) 5' and (c) 3' products of clone 1912-d05. The 1-kb marker from the original gel was superimposed onto the autoradiograph using the Gel-Pro computer program.

Therefore, the results suggested that the orientation of the I-1 gene is: the 3-kb fragment at the 5' end; the 2·7-kb band in the middle; and the 2-kb band at the 3' end (Fig. 3.7).

Figure 3.7. The preliminary orientation of the I-1 gene and restriction fragments predicted by hybridization of 1C1A DNA (type 1 clone) with radio-labelled 5' and 3' fragments of 1912-d05. The scale bar is shown.



The restriction maps were determined for clone types 1 and 2 (clones 1C1A and 4A1A, respectively) from the various restriction digests (Figs A5 & A6, Appendix 1, pp. 179 & 180, respectively). The predicted orientation of the I-1 gene was confirmed from the restriction map of 1C1A.

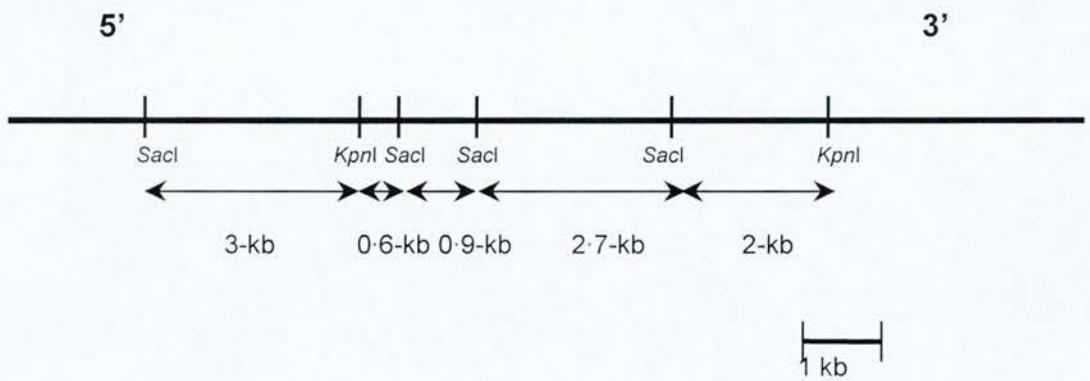
3.2.7. Subcloning of the inhibitor-1 gene

In order to facilitate sequencing of clone type 1, it was decided to subclone the positive restriction fragments determined by Southern blotting (i.e. hybridization with the 336-bp PCR fragment; Fig. 3.3) into a suitable vector. Therefore, the *Kpn*I/*Sac*I restriction fragments from clone type 1 (1C1A) of sizes 3, 2·7 and 2 kb were ligated into pBluescriptIIKS– (Fig. A1d, Appendix 1, p. 169). To complete the subcloning of clone type 1, the non-hybridizing 900-bp and 600-bp restriction fragments (determined by the restriction map of 1C1A as shown in Table 3.7) were also subcloned into pBluescriptIIKS–. To confirm overlap of the two clone types, the 2-kb *Hind*III fragment from each clone was subcloned into pBluescriptIIKS–.

3.2.8. The sequence of the mouse inhibitor-1 gene

The subcloned fragments from the genomic clones 1C1A and 4A1A were commercially sequenced by DNASHEF (Edinburgh) and MWG-Biotech (Germany) for reasons of cost and convenience since no cost was incurred if no sequence was obtained. In some instances, the same plasmid DNA was sent to different companies for sequencing (Fig. A4, Appendix 1, p. 176, for comparative analysis). A total of 9 kb of genomic DNA was sequenced (Fig. A7, Appendix 1, p. 181). Nucleotide 1 does not refer to the position of the transcription start site since the promoter regions are not known, but does refer to the 5' flanking region of DNA that was sequenced.

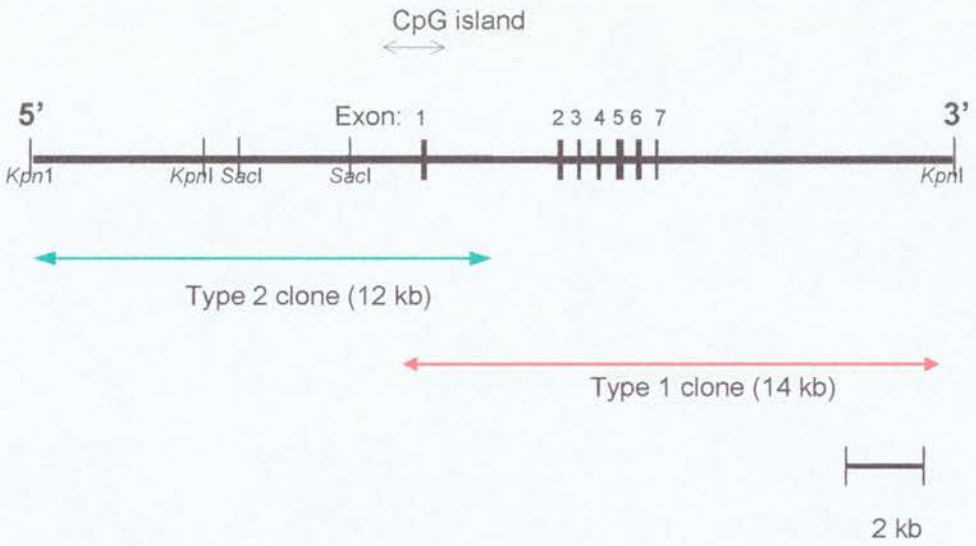
Figure 3.8. The final arrangement of subclones which contain the I-1 gene spanning ~9 kb of genomic DNA. The *Sac*I and *Kpn*I restriction sites and the scale bar are shown.



Comparison of the 2-kb *Hind*III fragments from the clone types 1 and 2 showed strong sequence homology (Fig. A8, Appendix 1, p. 186). Therefore, the two clones overlapped as predicted from the restriction maps and Southern blot analysis. From the restriction map, the order of the subclones from type 1 clone was then determined. Analysis of ~ 500 bp of 5' sequence (i.e. the first 300 bp immediately before the first exon, the first exon and 80 bp immediately after the first exon) has a CpG score of 0·74. This was derived from an algorithm based on work by Gardiner-Garden & Frommer (1987) to identify CpG islands (i.e. GC-rich regions) in the vertebrate genome. The algorithm has a scale between 0 and 1, with the values closer to 1 indicating the presence of a CpG island. These islands

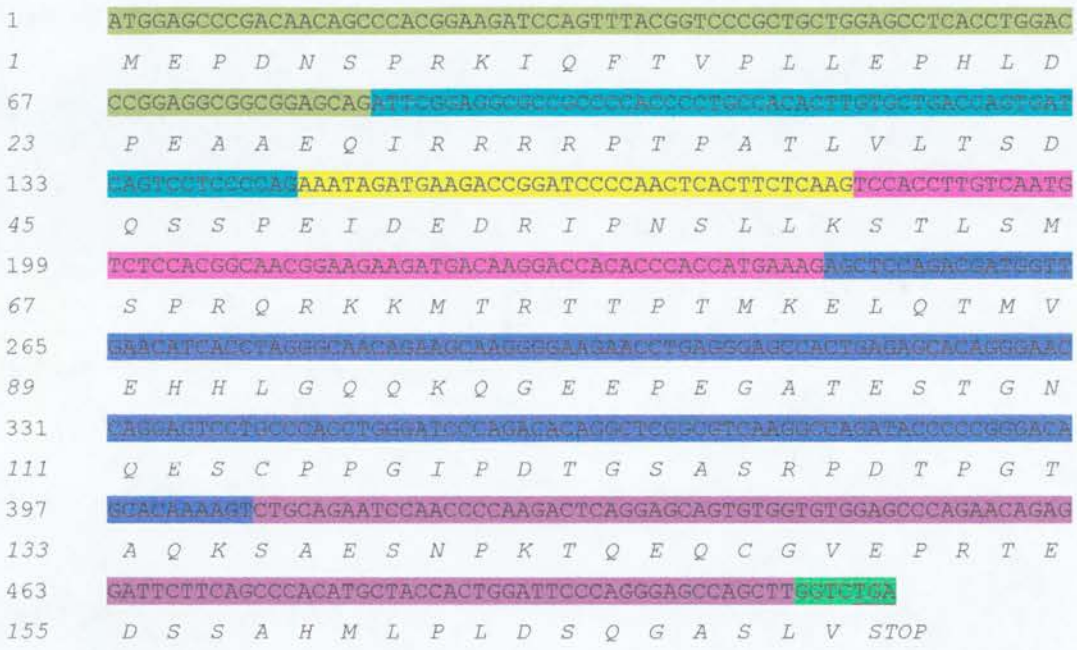
are often present in the 5' region of housekeeping and tissue-specific genes (although some genes have 3' CpG islands). The 5' CpG island may extend into the translation start site in protein-coding genes, which is the case for the I-1 sequence (Fig. 3.9).

Figure 3.9. Diagram showing the overlap of clones type 1 and 2 isolated from the library (represented by the red and green arrows, respectively). The seven coding exons for the predicted mouse I-1 protein span 7 kb of genomic DNA. The position of the CpG island, selected restriction sites and the scale bar are shown.



NIX analysis of the subcloned I-1 fragments showed that the encoding region for the predicted protein sequence spans a 7 kb region of genomic DNA and is encoded by seven exons. Furthermore, the mouse mRNA and protein sequences were predicted (Fig. 3.10) by comparison of the genomic sequences, the various 1912-d05 and 3141-o09 I.M.A.G.E. sequences, plus the remaining partially sequenced I.M.A.G.E. clones available from I.M.A.G.E. The computer program Translate (http://expasy.ch/tools/dna.html) was used to determine the protein sequence from the submitted mRNA sequence using the three possible alignments analysed for both sequence directions (i.e. 5' to 3' and 3' to 5'; p. xvi for the list of amino acids and their one letter code).

Figure 3.10. The coding sequence of mouse I-1 mRNA showing the position of the exons 1–7: Exon 1, exon 2, exon 3, exon 4, exon 5, exon 6 and exon 7 (including the stop codon, underlined). The corresponding predicted amino acid sequence is shown in italics beneath the mRNA sequence.



The predicted mouse I-1 sequence was compared to that of rat and there were a total of 25 nucleotide difference between the two sequences. From the deduced mRNA sequence, the predicted mouse protein sequence was determined and compared to the known I-1 sequences (Fig. 3.11). Thus, the mRNA sequence is 513 bp in length, and as predicted, is similar to the rat inhibitor-1 mRNA sequence, sharing 95·1% homology (Fig. 3.11).

Figure 3.11. Comparison of mouse, rat (Elbrecht *et al.* 1990), human (Endo *et al.* 1996) and rabbit (Aitken *et al.* 1982) protein sequences by Clustal analysis: (–) a space in the sequence to give best alignment; (*) 100% homology for all sequences. The stop codon is not included. The shared KIQF motif at amino acid sites 9–12, and the threonine at amino acid 35, are highlighted in blue and are important for PP-1 inhibition. The differences between the mouse and rat sequences are indicated by the amino acids highlighted in pink.

```
           1                                                          60
MOUSE      MEPDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEIDEDRIPNSLL
RAT        MEPDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEVDEDRIPNPLL
HUMAN      MEQDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEIDEDRIPNPHL
RABBIT     MEQDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEVDEDRIPNPLL
           ** ********************************************** *******  *
           61                                                        120
MOUSE      KSTLSMSPRQRKKMTRTTPTMKELQTMVEHHLGQQKQGEEPEGATESTGNQESCPPGIPD
RAT        KSTLSMSPRQRKKMTRTTPTMKELQTMVEHHLGQQKQGEEPEGATESTGNQESCPPGIPD
HUMAN      KSTLAMSPRQRKKMTRITPTMKELQMMVEHHLGQQQQGEEPEGAAESTGTQESRPPGIPD
RABBIT     KPSLAMSPRQRKKMTRTTPTMKELQMMVEHHLGQQEQGEEPEGAAEGTGAQESQPPGTPG
           *  * *********** ******** ********* ******** * ** *** *** *
           121                                                171
MOUSE      TGSASRPDTFGTAQKSAESNPKTQEQCGVEPRTEDSSAHMLPLDSQGASLV
RAT        TGSASRPDTSGTAQKPAESKPKTQEQRGVEPSTEDLSAHMLPLDSQGASLV
HUMAN      TEVESRLGTSGTAKKTAECIPKTHERGSKEPSTKEPSTHIPPLDSKGANSV
RABBIT     TGAESRLGPSATAQKPAQPSPRAQERRGEEPSTAKTS-Q----DSQGASAV
           *   **     ** * *   *   *    ** *   *      ** **  *
```

Comparison of the mouse protein inhibitor-1 sequence with rat, human and rabbit showed strongest homology in the amino region of the protein. The main differences in the amino acid sequence were toward the carboxyl end, a finding which is consistent with previous comparisons of the sequences (Endo *et al.* 1996). The predicted mouse I-1 sequence has a KIQF motif at amino acid sites 9–12 and a threonine at residue 35, both of which have been shown to be important for inhibition of PP-1 (Endo *et al.* 1996).

3.2.9. Chromosomal localization of the inhibitor-1 gene

As a final characterization of the mouse I-1 gene, the location of the I-1 gene was determined. The 14-kb clone type 1 isolated from the library was an adequate size to use for FISH analysis. Ms Shelagh Boyle determined the chromosomal localization of the gene since she has the expertise and facilities for the procedure. The I-1 gene was located at the end of the long arm of chromosome 15, band F. Figure 3.12 shows that the I-1 FITC signal is apparent on each chromatid of chromosome 15 and there is a positive signal in an adjacent interphase cell. A more detailed localization of the gene was not relevant to this project.

According to the Chromosome 22 Mapping Group at the Sanger Centre (http://www.sanger.ac.uk/HGP/Chr22/), there is synteny for human chromosome 22 and mouse chromosome 15. Other genes located in this area of mouse chromosome 15 include *Wnt-1* and the *Hox C* cluster. A transgene-induced deletion that causes perinatal lethality (ple) has also been mapped to 15F (Beier *et al.* 1989). A number of known mouse mutants map to this region. These include shaven (Sha), crimpy (cpy), Naked (N) and caracul (Ca), which affect skin and hair texture, and probably involve mutations of the Keratin type II complex [Compton *et al.*, 1991; data from Blake *et al.* (2001) at the Mouse Genome Database (MGD), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine USA; April 2001; http://www.informatics.jax.org/].

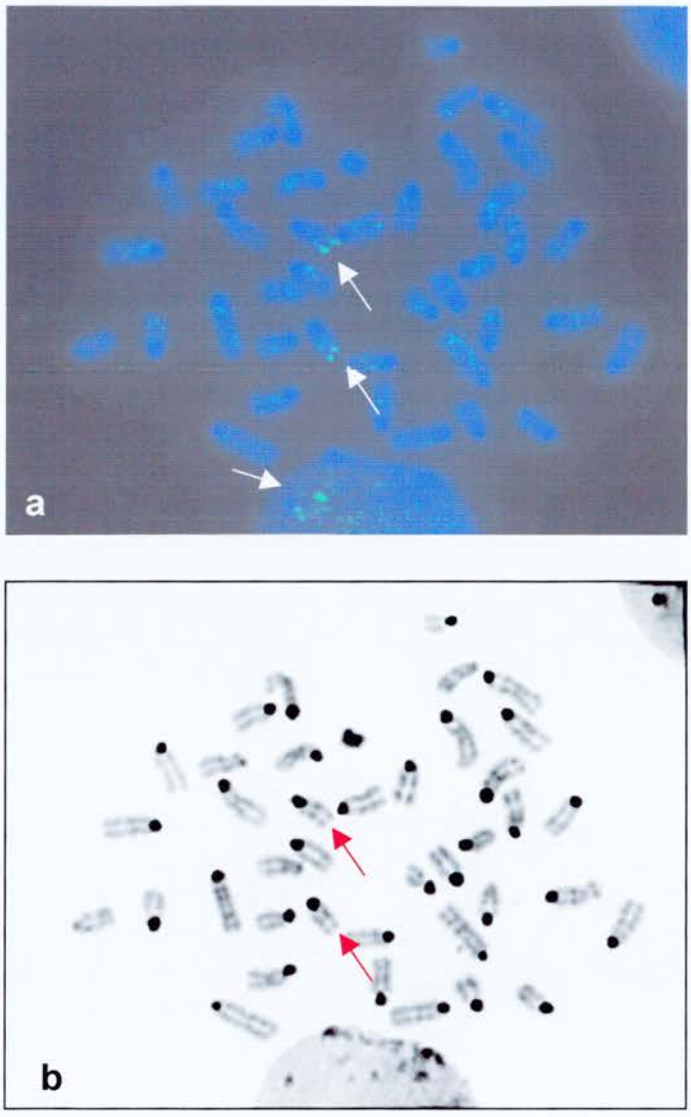


Figure 3.12. Chromosomal localization of the I-1 gene on normal mouse metaphase spreads using fluorescent *in situ* hybridization. (a) The white arrows show the positive signal on the metaphase spread and in a nearby interphase cell. (b) The Giemsa-banded chromosomes show that the gene is located on chromosome 15F, as indicated by the red arrows.

3.3. Discussion

The aim of this research chapter was to screen a mouse genomic library and isolate clones which contained homology to the rat I-1 mRNA sequence, which could then be used to produce transgenic mice. The screening of the library resulted in two types of clone being isolated which contained such homology. This allowed the mouse I-1 gene to be successfully characterized by restriction digestion, sequence analysis and FISH.

3.3.1. I.M.A.G.E. clone analysis and library screening

The initial section of this research chapter confirmed that the correct I.M.A.G.E. cDNA clones were sent from HGMP-RC, which was important to establish since one of the clones was selected to screen the mouse genomic library. The 5' 336-bp product obtained by PCR from the 1912-d05 clone was used in preference to a restriction digest fragment. This was because the size of the 5' fragment could be accurately predicted and there was no guarantee that a suitably sized fragment could be obtained by restriction digest: this was confirmed by sequence analysis with the Webcutter computer program. Fortunately, there was a *Bam*HI site within the 1912-d05 sequence that produced a 3' fragment of suitable length (~280 bp).

Screening of the λ129 mouse genomic library isolated clones containing homology to the rat I-1 mRNA sequence. The automatic subcloning facility of the library allowed easier handling and manipulation of the resulting isolated clones. Restriction digestion was used to analyse 18 clones isolated from the library and this showed the presence of four different types of clone. PCR analysis of these clones was less accurate for characterization since only three types of clones were identified: the presence of introns in the genomic sequence is the likely cause of this.

Southern blot analysis confirmed that there were two different types of clones containing I-1 sequence homology (termed type 1 and type 2 with inserts of ~14 and ~12 kb, respectively). Fifteen out of the 18 clones were positive for I-1

sequence homology: the type 1 clone was the predominant recombinant isolated from the library (13 out of these 15 I-1-positive clones were of type 1). The low number of negatives was probably a result of the use of the 5' PCR product for screening since presence of the 3' homopolymeric region would have had homology with other mRNAs, and negative clones would have then been selected. Furthermore, only well-isolated plaques were picked after the secondary screening which reduced the selection of negative clones.

3.3.2. Characterization of the protein-coding region of the inhibitor-1 gene

The restriction maps of clone types 1 and 2 were determined by single, double and triple digests of the DNA with a range of enzymes. Subcloning of the positive fragments (identified by Southern blotting) from the two clones was successful. The two remaining non-hybridizing fragments to rat I-1 mRNA from the type 1 clone (600 and 900 bp in size) were also subcloned, which allowed ~9 kb of genomic DNA to be sequenced (Fig. 3.6 and Fig. A7, Appendix 1, p. 181). Analysis of these subcloned sequences by the NIX computer program showed that clone type 2 contained the first exon and ~9 kb of flanking upstream 5' sequence, most of which remains unsequenced. The clone type 1 contained the seven I-1 protein-coding exons. Therefore, there was overlap between types 1 and 2 clones since both contained the first protein coding exon.

The predicted mRNA sequence for the mouse I-1 protein was obtained by comparison of: the 10 partially sequenced I.M.A.G.E. clones; genomic sequences; and the various commercial sequences of clone 1912-d05. This resulted in a reliable final mRNA sequence. The mRNA coding sequence was 513 bp in length and was strongly homologous to rat I-1, sharing 95·1% identity; and the deduced I-1 protein sequence contains 171 amino acids, sharing 95·3% homology to rat. There were a total of 25 nucleotide differences between the mouse and rat mRNA sequences, resulting in eight amino acid differences. Only one of the amino acid changes was conservative (amino acid 50; isoleucine and valine in the mouse and rat sequences, respectively): the remaining seven differences were non-

conservative. Comparisons of the mouse protein I-1 sequence with rat, human and rabbit showed strongest homology in the amino region of the protein. The main differences in the amino acid sequence were toward the carboxyl end, which is consistent with previous comparisons of the sequences (Endo *et al.* 1996; McLaren *et al.* 2000, Appendix 2, p. 188).

Two elements in the I-1 protein structure which are required for efficient PP-1 inhibition are a KIQF motif at the amino end of the sequence and a threonine at position 35 (Endo *et al.* 1996). Both features are present in the mouse, rat, rabbit and human inhibitor-1 sequences. Aitken *et al.* (1982) concluded that phosphorylation of the serine at amino acid site 67 (Ser^{67}) does not regulate I-1 function. However, Huang & Paudel (2000) used recombinant human I-1 to show that phosphorylation of Ser^{67} does activate I-1, possibly in a similar mechanism to Thr^{35} phosphorylation.

CpG islands contain non-methylated CpG dinucleotides and are usually associated with coding regions of genes. The CpG score for the mouse I-1 gene is 0·74, using a scale between 0 and 1 based on work by Gardiner-Garden & Frommer (1987): the values closer to 1 indicate the presence of a CpG island. These islands are often present in the 5' region of housekeeping and tissue-specific genes, and may extend into the translation start site, which is the case for the mouse I-1 gene: the 5' CpG island lies within a region extending 250 bp before to 70 bp past the first exon. As a final characterization of the mouse I-1 gene, FISH was used to localize the gene to chromosome 15F and there may be synteny with human chromosome 22. No mutant phenotypes map to the region for which I-1 may be a candidate gene. As it is now known that *I-1*$^{-/-}$ mice have no overt phenotype (Allen *et al.* 2000), this is not surprising.

The work carried out in this research chapter confirmed the strong homology of the I-1 gene across a range of species. A total of 9 kb of genomic DNA was sequenced, within which lay the seven protein-coding exons (Fig. A7, Appendix 1, p. 181). The type 2 clone isolated from the library had ~9 kb of 5' upstream

sequence that could be used to build expression constructs with a view to characterizing the mouse I-1 promoter regions (Chapter 5).

Chapter 4

Inhibitor-1 protein expression in the developing mouse embryo

4.1. Introduction

4.1.1. Abstract

The mouse I-1 protein expression pattern in the developing kidney was analysed using whole-mount tissue, a commercially available antibody for I-1 and confocal microscopy. The initial results showed that I-1 was only expressed in the peripheral layer of cells in mouse kidneys from E12·5 through to adult. No other cells in the kidney expressed I-1. Analysis of early kidneys showed I-1 expression in the coelomic epithelium of embryos at stage ~E11. The phenotype of the I-1 positive cells suggested that they were epithelial, and this was confirmed using antibodies against laminin and pan cytokeratin. The expression pattern suggested that the I-1 antibody labelled the mesothelium, a cellular layer that bounds the internal coelomic cavity. This was confirmed by I-1 labelling of tissue from the coelom (i.e. liver, lung, heart, gonad and intestine). Inhibitor-1 was also present in the surface layer of ectoderm in the developing embryo and adult mouse.

4.1.2. Aim of the chapter

Chapter 3 determined the genomic sequence of the I-1 gene, and from this, the predicted mRNA and protein sequences. In order to further characterize the mouse I-1 gene, its protein expression pattern in the developing mouse kidney was analysed. This would determine whether I-1 is a kidney stem cell marker, as suggested by Svennilson *et al.* (1995). Although cell type-specific markers have been described for other cellular compartments of the developing kidney (e.g. BF-2 for stromal cells and Pax-2 for the stem cells, the condensed tubules and the ureteric bud epithelium), there is no unique marker for the kidney stem cells. Therefore, whole-mount tissue, a commercially available antibody and confocal microscopy were used to identify the I-1 protein expression pattern in the developing mouse kidney.

4.1.3. Background and experimental approach

The use of whole-mount tissue for the analysis of protein expression eliminates the need for processes such as wax embedding, which can introduce factors, such as non-specific fluorescence as a result of contaminated solutions. Also, there is less handling of the specimen. Analysis of whole-mount tissue by confocal microscopy has the advantage that the optically sectioned tissue can be observed in terms of the tissue or embryo as a whole. This is not possible in sectioned material, unless images of the tissue are layered using computer graphics to produce a three-dimensional image. Therefore, whole-mount kidney tissue, a commercially available I-1 antibody, and confocal microscopy were used to analyse the I-1 expression pattern in the developing kidney. The expression pattern was compared to the I-1 expression pattern in kidney sections.

If I-1 is a kidney stem cell marker, as predicted by Svennilson *et al.* (1995), then it would be expressed in the uncondensed, unepithelialized cells at and near the periphery of the kidney. These cells would express Pax-2, a marker of all stages of nephrogenic mesenchyme (Dressler *et al.* 1990), but they would not yet have epithelialized and so have no basal lamina and, thus, would not express laminin. Therefore, antibodies against I-1, Pax-2 and laminin proteins could be used to characterize the stem cells of the kidney.

The only published data of I-1 protein detection on kidney sections was by Fryckstedt *et al.* (1993) who analysed the I-1 and DARPP-32 protein expression patterns using sectioned rat kidneys. There is no published analysis of I-1 protein expression using whole-mount tissue and confocal microscopy. The commercially available polyclonal sheep anti-rabbit I-1 antibody (Serotec) was obtained, which was based on the work by MacDougall *et al.* (1989), who detected I-1 in immunoblots using antibodies from both the phosphorylated and dephosphorylated I-1 protein as immunogens. Thus, the phosphorylated state of the protein used to obtain the antibody had no effect on I-1 detection. An anti-sheep FITC-labelled secondary antibody (Sigma) was used to visualize the I-1 antibody.

4.2. Results

4.2.1 Expression of inhibitor-1 in the mouse whole-mount kidney

Analysis of I-1 expression in mouse kidneys from a range of ages (E12·5 through to adult) showed that the protein was only detected in the cytoplasm of the outermost layer of cells in all of the ages studied. Figures 4.1.a–c (E12·5, E13·5 and E16·5 kidneys, respectively) showed that the I-1 protein was localized in the surface, peripheral cellular layer. There was no I-1 staining within the kidney. The morphology of the I-1-positive cells suggested they were unilaminar, squamous epithelia (Williams *et al.* 1995). The I-1 antibody did not label within the kidney tissue and this suggested that I-1 was not localized to the stem cells of the MM.

E11 uninduced MM and E11·5 induced MM (including presence of the coelomic cavity) showed I-1 expression in the coelomic epithelium, which contains the developing kidney (Fig. 4.1.d). However, removal of the coelomic tissue and incubation of only the rudimentary kidney with the I-1 antibody showed no I-1 expression (Fig. 4.1.e). This result confirmed that the I-1 expression pattern was not an 'edge effect': if this was the case, then the positive peripheral staining would have been present after removal of the coelomic epithelium.

Kidneys incubated with serum, but no primary antibody, showed no I-1 expression in the tissue (Fig. 4.1.f). The use of PBS instead of PBSTX during the incubation steps did not alter the I-1 staining pattern although the background was slightly lower in those samples that had been processed in the presence of the detergent. Kidneys fixed with 4% PFA plus glutaraldehyde were analysed since Lowenstein *et al.* (1995) found that inclusion of this fixative gave stronger immunoreactivity, presumably because of better preservation of antigenic sites and/or retention of antigen. The presence of glutaraldehyde did not enhance the signal nor reveal other cell types expressing I-1 in the kidneys. Also, methanol-fixed kidneys did not show I-1 expression, suggesting that the antigen was lost or masked using this fixative.

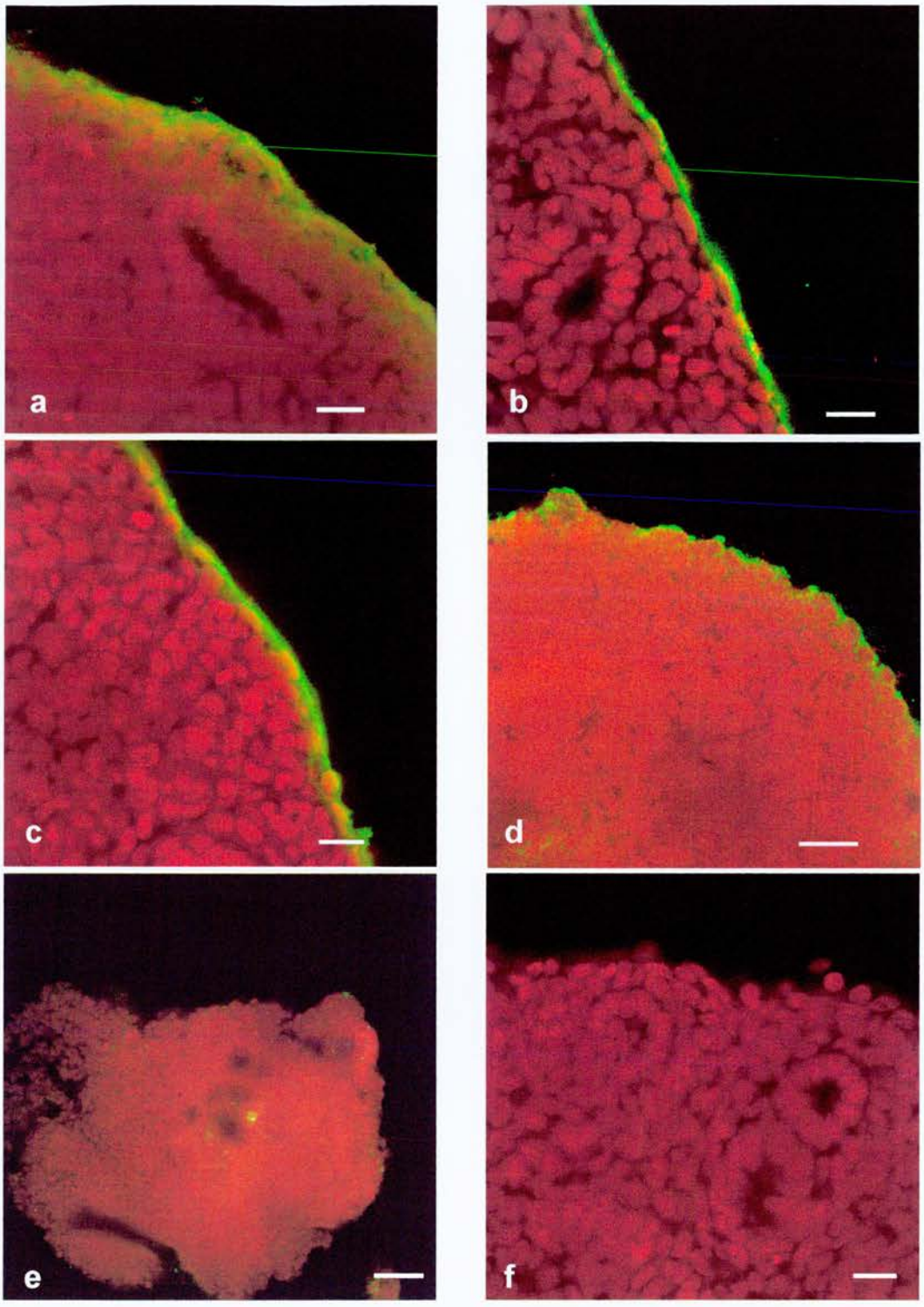


Figure 4.1. Inhibitor-1 expression in the developing mouse kidney. The cells were counter-stained with propidium iodide (red). The inhibitor-1 protein (FITC) is expressed by the peripheral layer of cells in: (a) E12·5; (b) E15·5; (c) E16·5; and (d) E11·5 kidneys. (e) The E11·5 kidney dissected free of the coelomic epithelium does not express inhibitor-1. (f) Kidneys incubated without the I-1 antibody do not give a positive signal. Bars: (a – d & e) 20 μm; and (f) 65 μm.

4.2.2. Expression of inhibitor-1 in wax embedded kidney sections

To compare the I-1 expression pattern of the whole-mount kidneys with sectioned material, kidney sections were incubated with the I-1 antibody. E15·5, E16·5 and newborn kidneys were fixed in 4% PFA, dehydrated through alcohol and xylene, and then embedded in paraffin wax. Following this, 7-μm sections were placed onto TESPA slides. The sections were incubated with the I-1 antibody, as described in Chapter 2 ('Materials and methods'), and visualized using the FITC-labelled secondary antibody, although no propidium iodide was added as a counter-stain. As a control, no I-1 antibody was applied to the section. The results showed that there was a positive signal around the kidney periphery (Figs 4.2.a–c), similar to the whole-mount kidneys incubated with the I-1 antibody (Figs 4.1.a–d and section 4.2.1, above). However, all sections (including those incubated without the I-1 antibody) had high levels of non-specific background and the occasional highly fluorescent cell (Figs 4.2.d–f). This observation was similar to the results by Fryckstedt *et al.* (1993).

The whole-mount kidneys incubated with the I-1 antibody gave very low background in comparison to the kidney sections and the resolution of the tissue was better than sectioned material. Therefore, whole-mount tissue was used for future analyses.

4.2.3. Expression pattern of Pax-2 protein in the mouse whole-mount kidney

The anti-Pax-2 antibody identifies the nuclei of condensing mesenchyme-to-epithelial cells and the duct tubules of the kidney (Dressler & Douglass 1992). If there is expression of I-1 in the metanephric mesenchyme, then it would be expected to co-express with Pax-2. E13·5–15·5 whole-mount kidneys were incubated with the antibody using the protocol described in Chapter 2. The condensing metanephric mesenchyme was labelled with the Pax-2 antibody (Figs 4.3.a–c). Omission of the Pax-2 antibody gave no nuclear staining (4.3.d).

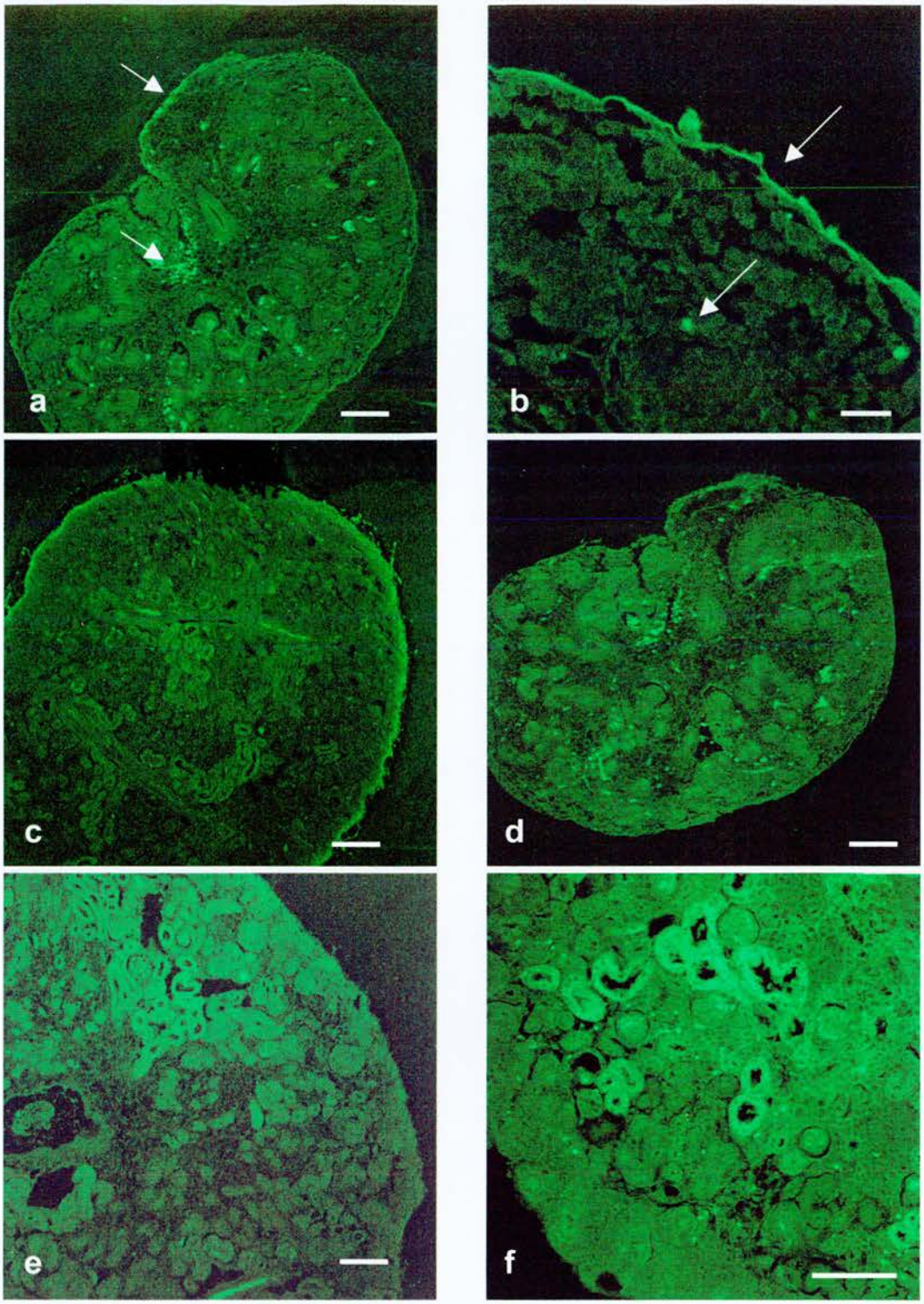


Figure 4.2. Inhibitor-1 expression in wax-embedded kidney sections. The inhibitor-1 protein was visualised with FITC but no counter-stain was used. (a & b) E15·5 and (c) E16·5 kidneys show inhibitor-1 expression in the peripheral layer of cells and some internally labelled cells (a & b, white arrows). (d) E15·5, (e) E16·5 and (f) newborn sections incubated with no inhibitor-1 antibody show no peripheral labelling, but the internal cells give a positive FITC signal. Bars: (1, 3 & 4) 100 μm; (2) 20 μm; and (5 & 6) 50 μm.

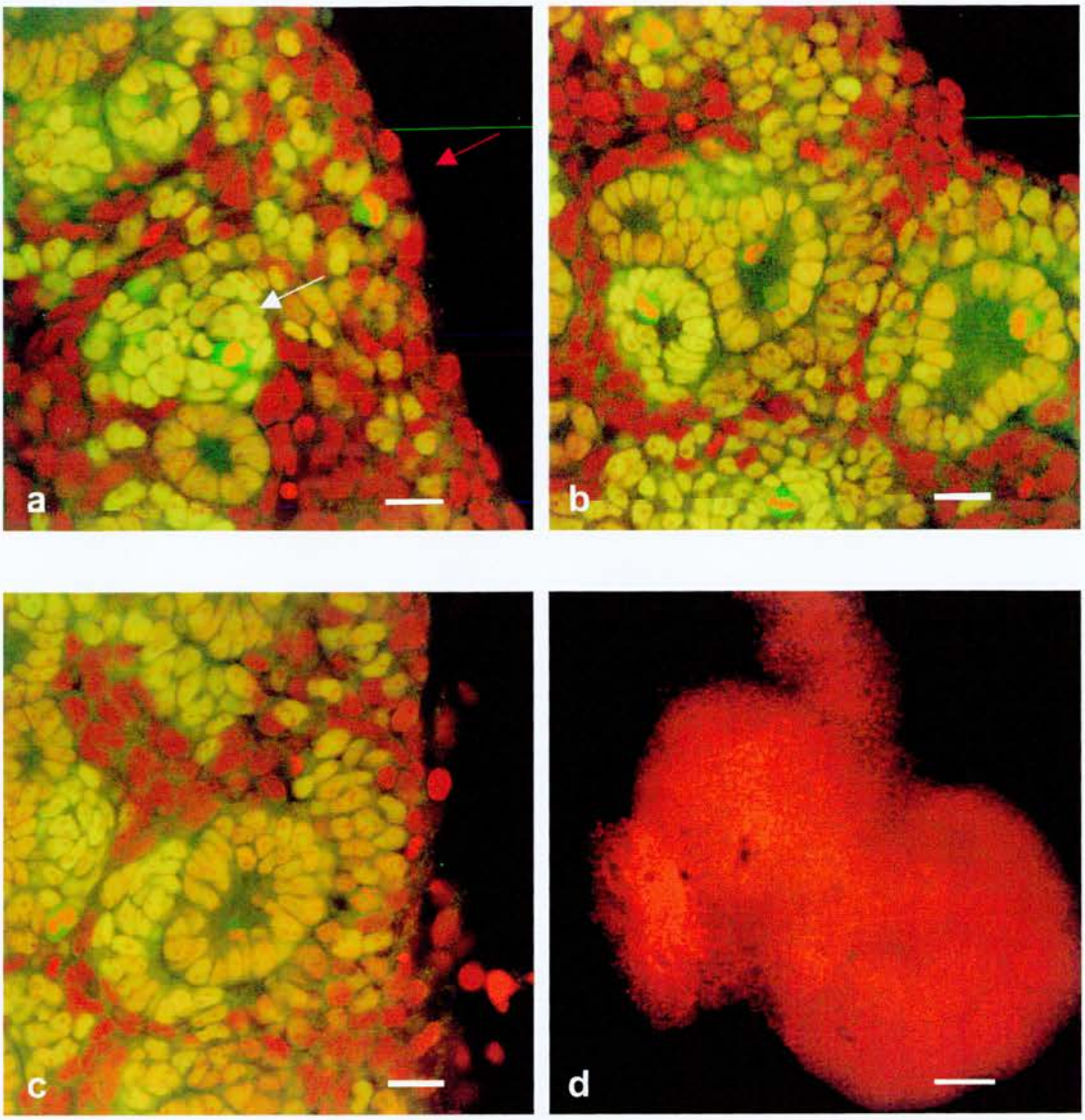


Figure 4.3. Pax-2 expression in whole-mount (a) E13·5, (b) E14·5 and (c) E15·5 mouse embryonic kidneys. (d) The control E13·5 kidney was incubated in the absence of the Pax-2 antibody. The condensing mesenchyme is positive for Pax-2 (FITC, indicated by the white arrow) but the peripheral layer of cells that are positive for I-1 (Fig. 4.2) are negative for Pax2 (red arrow). Bars: (a – c) 20 μm; and (d) 100 μm.

Double labelling of kidneys using the Pax-2 and I-1 antibodies were not successful because of the difference in incubation conditions required for the antibodies.

Bard *et al.* (1996) referred to stem cells as early condensates, and therefore, these cells would be identified by the Pax-2 antibody. However, there was no co-expression of Pax-2 and I-1 in the E13·5–E15·5 kidneys analysed (Fig. 4.3.a) suggesting that I-1 is not a marker of induced metanephric mesenchyme, and hence, I-1 is not a stem cell marker in the developing kidney.

4.2.4. Expression of laminin in the mouse whole-mount kidney

The I-1-positive cells determined in section 4.2.1 had unilaminar epithelium morphology, according to classification by Williams *et al.* (1995). To test the hypothesis that these I-1-positive cells had an epithelial phenotype, whole-mount kidneys were incubated with an anti-laminin antibody. Laminins are a major component of basement membranes (Chapter 1, section 1.5.3). In all of the kidneys analysed (E13·5–E16·5), nephrogenic and duct tubules had basal laminae and these provided internal controls for examining the peripheral cells. Figure 4.4.a shows weak and incomplete staining of the basal lamina in the underlying single layer of cells encapsulating the E12·5 organ. However, expression levels increased by E13·5 (Fig. 4.4.b) and continued to do so as the kidney developed (Fig. 4.4.c showing an E16·5 kidney). Kidneys incubated without the laminin antibody did not show any laminin staining (Fig. 4.4.d).

The laminin expression observations were compatible with the I-1-positive cells having an epithelial phenotype, and overall, the results suggested that the I-1 protein was expressed by the mesothelium, a cellular layer that covers the tissue of the coelomic cavity and is derived from the coelomic epithelium. However, no internal epithelia (i.e. endodermal cells) expressed I-1, and thus, I-1 was not a general epithelial marker.

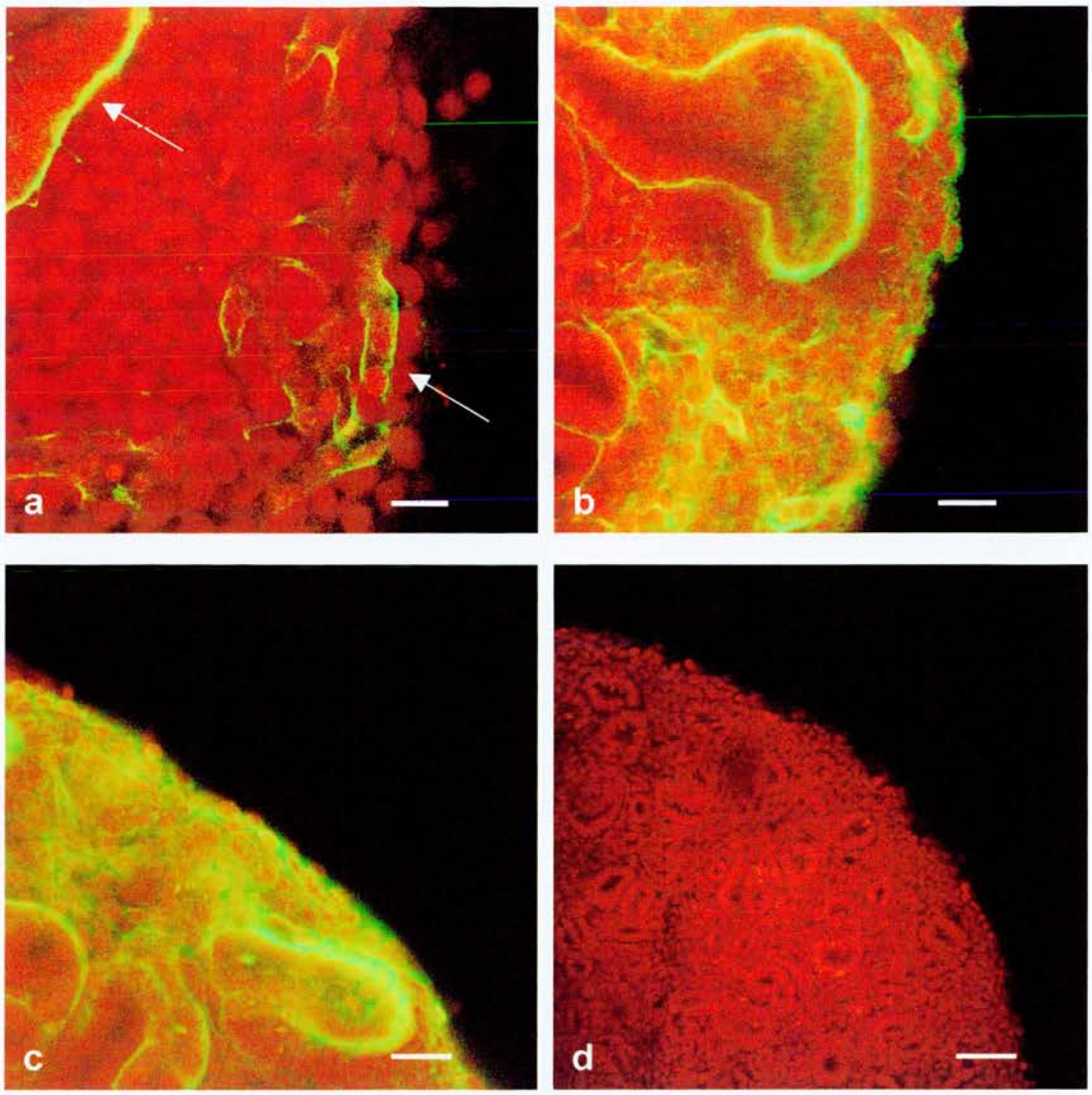


Figure 4.4. Laminin expression in the developing mouse kidney. Laminin was visualized with FITC and the tissue was counter-stained with propidium iodide. The internal basal laminae expressed laminin but the surface basal lamina was weak in (a) E12·5 and (b) E13·5 kidneys. (c) Laminin expression did increase as the kidney aged (E16·5). (d) E14·5 kidney incubated with no laminin antibody showed no laminin expression. Bars: (a & b) 20 μm; (c) 50 μm; and (d) 100 μm.

4.2.5. Expression of pan cyokeratin in the mouse whole-mount kidney

To confirm the laminin expression pattern, an anti-pan cyokeratin antibody was used to analyse E13·5 kidneys. Methanol-fixed kidneys were incubated with the antibody (Figs 4.5.a & b). The internal epithelial cells were strongly labelled by the antibody, but the peripheral layer of epithelial cells expressed the protein more weakly. This observation was similar to the laminin expression pattern. These results confirmed that the I-1-positive peripheral cells were epithelial in phenotype. Kidneys which had been fixed using 4% PFA were negative, and thus, the cytokeratin protein was lost using this fixative. Kidneys incubated in the absence of the primary antibody did not express the protein as a result of loss or masking of the antigen.

4.2.6. Summary of kidney analysis

Analysis of the developing mouse kidney suggested that the I-1 protein was expressed only in the outermost peripheral layer of cells. These cells were not metanephric mesenchyme, as confirmed by the Pax-2 protein expression pattern. The laminin and pan cyokeratin results showed that the cells had an epithelial phenotype, although these epithelial markers were weakly expressed in early kidneys. Therefore, these results suggested that I-1 expression was restricted to the mesothelium of the kidney in the developing mouse embryo. The mesoderm-derived mesothelium originates from the coelomic epithelium and lines the viscera of the body cavity (parietal mesothelium; Kaufman & Bard 1999).

4.2.7. Expression of inhibitor-1 and laminin in other mouse viscera

The I-1 expression pattern of the developing kidney suggested that the protein was restricted to the mesothelium. To test this new hypothesis, I-1 expression in other tissues of the developing mouse embryo was analysed (Table 4.1). Confirmation of the epithelial phenotype was obtained by incubation of mouse tissue with the laminin antibody (Table 4.2).

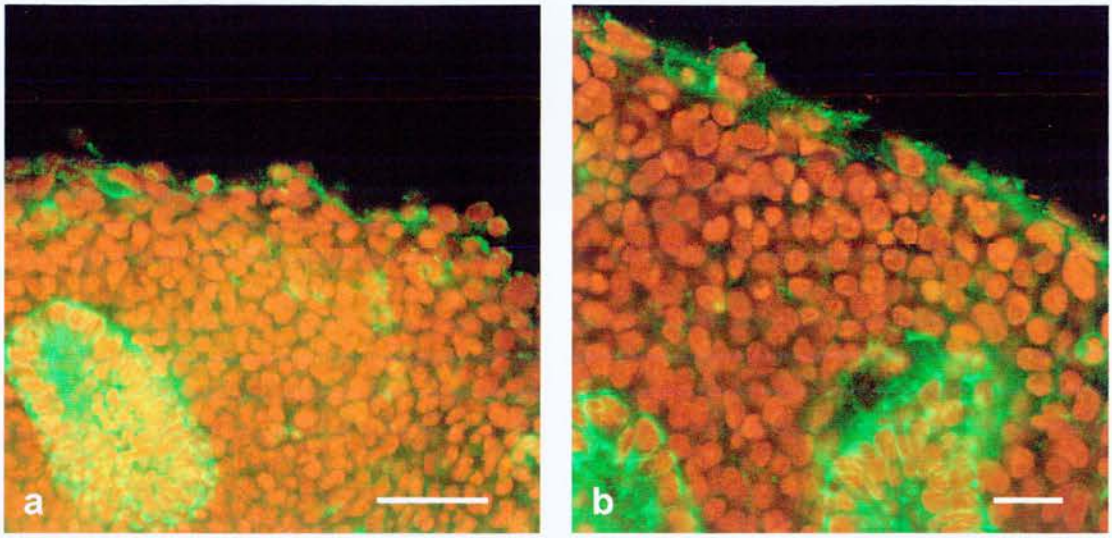


Figure 4.5. (a & b) Pan cytokeratin expression in methanol-fixed whole-mount E13·5 kidneys. Pan cytokeratin was visualized with FITC and the tissue counter-stained with propidium iodide. The epithelial cells were positive for pan cytokeratin: the peripheral layer of epithelial cells was more weakly labelled with the antibody than the internal epithelia, similar to laminin expression in the kidney (Fig. 5.4). Bars: (a) 40 μm; and (b) 20 μm.

	Embryonic age of mouse tissue (days)*										
Tissue	**9·5†**	**10·5†**	**11·5†**	**12·5**	**13·5**	**14·5**	**15·5**	**16·5**	**17·5**	**18·5**	**Adult**
Kidney	ND	ND	+	+	+	+	+	+	+	+	+
Lung	ND	ND	ND	+	+	+	+	+	+	+	+
Liver	ND	ND	ND	+	+	+	+	+	+	+	+
Heart	ND	ND	ND	+	+	+	+	+	+	+	+
Intestine	ND	ND	ND	+	ND	ND	+	+	+	+	+
Gonad	ND	ND	ND	+	+	+	+	+	+	+	+
Skin	+	+	+	+	+	+	+	+	+	+	+
Eye	+	+	+	ND	ND	ND	ND	+	ND	ND	ND

*Key: (+) positive staining for I-1; and (ND) not determined.
†Wholemount embryo used.

Table 4.1. Analysis of mouse tissue for inhibitor-1 (I-1) expression.

	Embryonic age of mouse tissue (days)*										
Tissue	**9·5†**	**10·5†**	**11·5†**	**12·5**	**13·5**	**14·5**	**15·5**	**16·5**	**17·5**	**18·5**	**Adult**
Kidney	ND	ND	+	+	+	+	+	+	ND	+	+
Lung	ND	ND	ND	ND	+	+	+	+	+	ND	+
Liver	ND	ND	ND	+	+	+	+	ND	+	+	+
Heart	ND	ND	ND	+	+	ND	+	+	+	+	ND
Intestine	ND	ND	ND	+	ND	+	+	+	+	ND	+
Gonad	ND	ND	ND	+	+	ND	+	ND	+	+	+
Skin	ND	ND	ND	ND	+	ND	ND	+	+	ND	+
Eye	ND	ND	ND	ND	ND	ND	ND	+	ND	ND	ND

*Key: (+) positive staining for laminin; and (ND) not determined.
†Wholemount embryo used.

Table 4.2. Analysis of mouse tissue for laminin expression.

Comparison of embryonic mouse lung tissue incubated with laminin and I-1 (Figs 4.6.a & b) showed that the peripheral layer of cells were I-1 positive and had an underlying basal lamina; therefore, they had an epithelial phenotype. Also, the basal lamina of lung and the other viscera strongly expressed laminin, unlike the kidney. Incubation of a range of tissue from various tissues showed that I-1 expression was limited to the single outer layer of surface epithelium of organs of the coelomic cavity (i.e. lung, intestine, heart, liver and gonad; Table 4.1 and Figs 4.6.b–f).

In the adult mouse, I-1 was also present in the surface epithelium of kidney, heart, lung, liver and the skin. There was no non-specific staining present in any of the tissues analysed (Table 4.1).

All basal laminae of the tissues analysed were labelled using the laminin antibody, showing that laminin is a general epithelial marker. When compared to the I-1 expression pattern, there was confirmation that I-1 is not a general epithelial marker: no internal epithelia were labelled and the I-1 protein was restricted to the peripheral epithelial layer, i.e. the mesothelium (Figs 4.1, 4.4 and 4.6.a & b).

4.2.8. Expression of inhibitor-1 during testis development

Analysis of embryonic gonads revealed I-1 expression of the mesothelium (i.e. the epithelial layer surrounding the gonads) at ~E12·5 (Fig. 4.7.b). At this stage of development, both male and female gonads are morphologically indistinguishable, and hence, are termed indifferent. However, I-1 was expressed by cells within a small number of the testes at ~E13·5 (Figs 4.7.c & d, showing presence of cells expressing I-1 within the testes). From E13·5 onwards, I-1 was expressed only in the mesothelial epithelium around the developing testes and not within the tissue (Fig. 4.6.f). This was consistent with developmental analysis describing mesothelial cell migration into the gonads in early testes formation (Buehr *et al.* 1993; Martineau *et al.* 1997; Capel *et al.* 1999; Kaufman & Bard 1999). These results also confirmed that the particular expression pattern for I-1 was not as a result of lack of penetration of the antibody within the tissue, since the internal

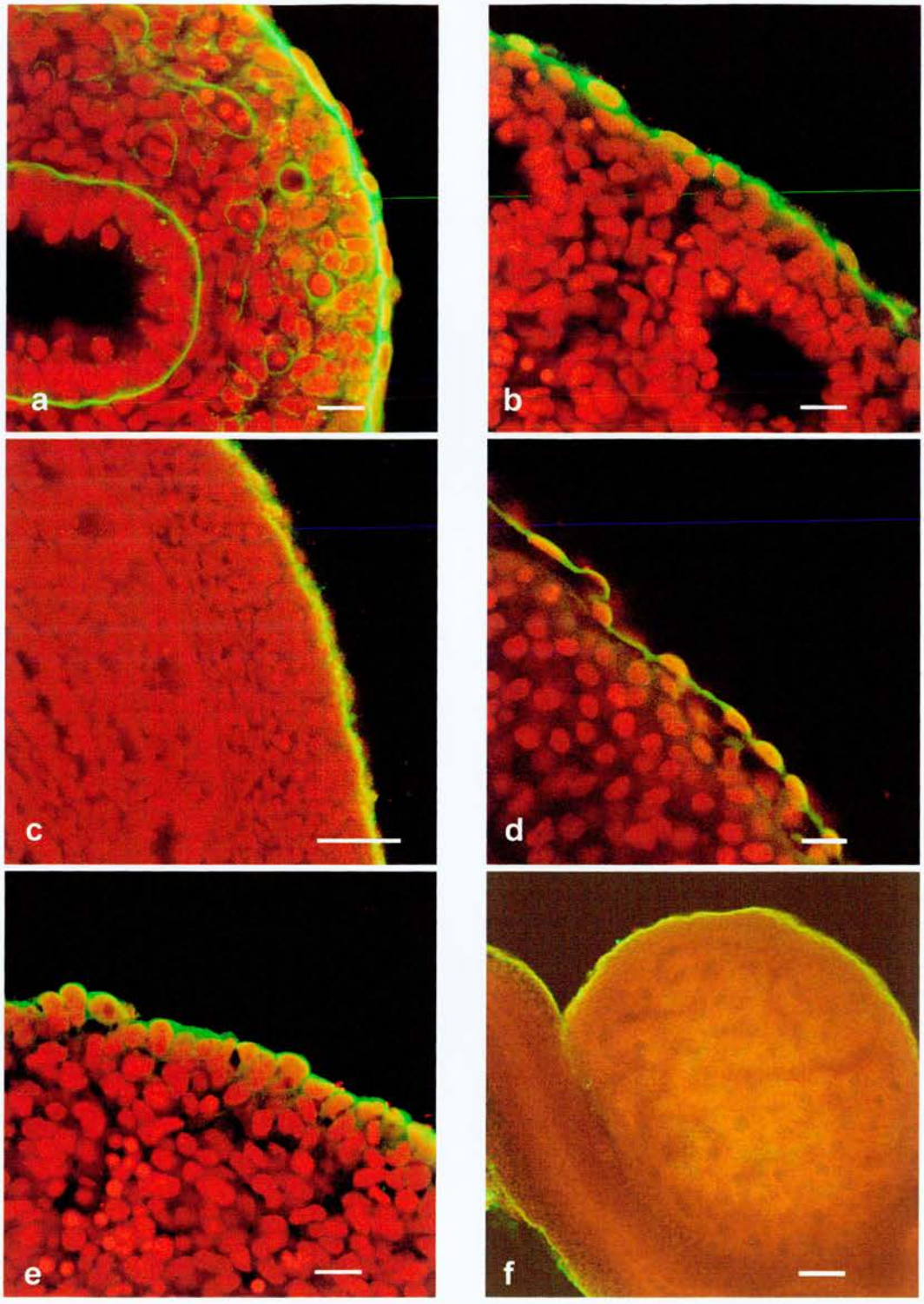


Figure 4.6. Laminin and inhibitor-1 expression in whole-mount mouse tissue. (a) Laminin expression in E12·5 lung. Inhibitor-1 expression in (b) E13·5 lung; (c) E17·5 intestine; (d) E12·5 heart; (e) E12·5 liver; and (f) E17·5 ovary. Bars: (a & b, d & e,) 20 μm; (c) 40 μm; and (f) 100 μm.

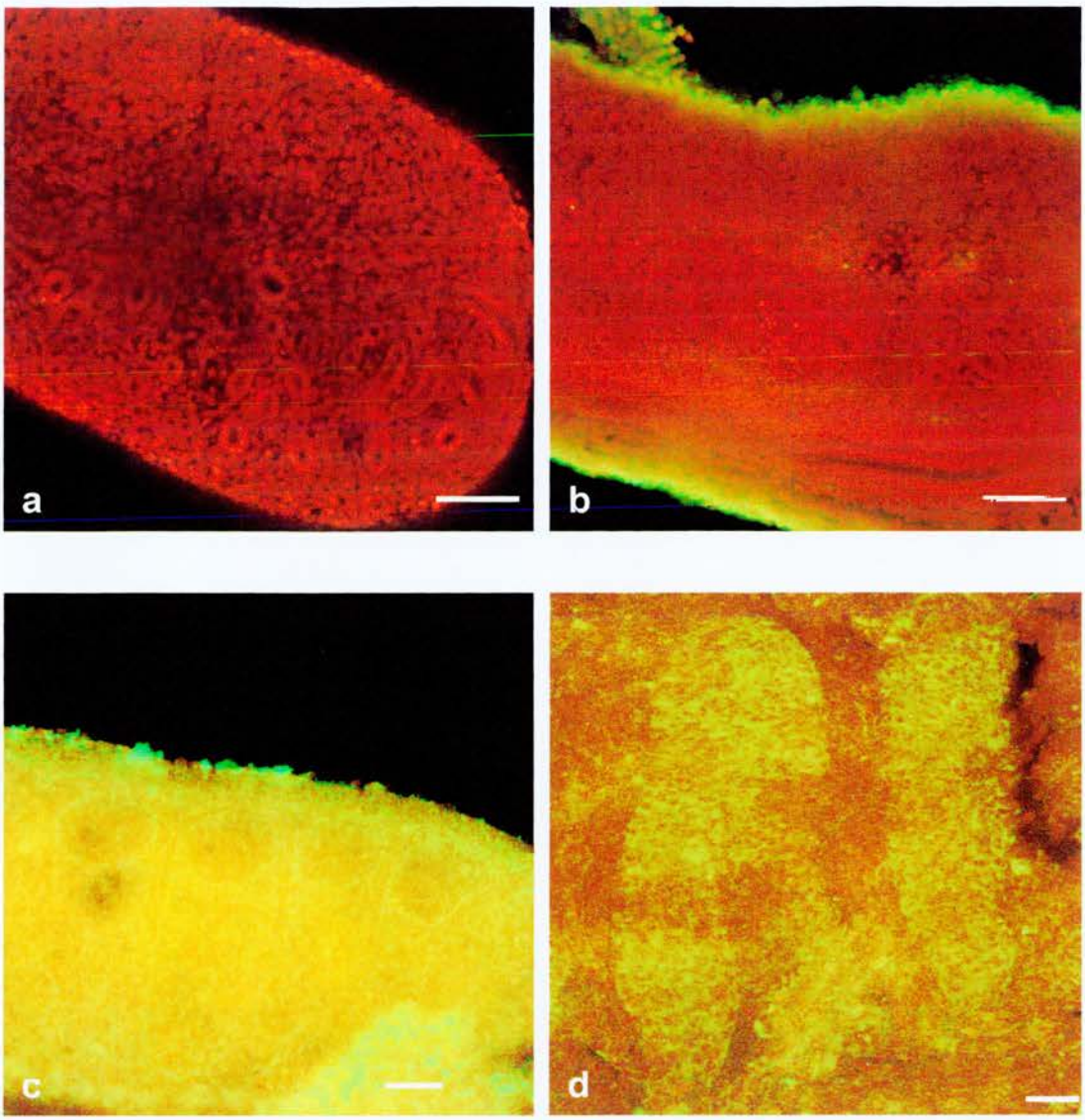


Figure 4.7. Inhibitor-1 expression in whole-mount mouse gonads. Inhibitor-1 was visualized with FITC and the tissue was counter-stained with propidium iodide. (a) Incubation of E12·5 gonad with no I-1 antibody gave no FITC labelling. (b) The E12·5 gonad labelled with I-1 showed positive I-1 labelling. (c & d) The sex cords of ~E13·5 testis were positive for I-1 expression, an observation only found in some of the E13·5 testes analysed. Bars: (a – c) 100 μm; and (d) 50 μm.

mesothelial cells were labelled. E12·5 gonads incubated without I-1 antibody did not show any staining (Fig. 4.7.a). Laminin expression of testes and ovaries confirmed the peripheral basal lamina underlying the I-1-positive layer of cells.

4.2.9. Expression of inhibitor-1 in other mouse tissue

There was I-1 expression in the outer epithelium or cornea of the developing eye, which is ectoderm-derived (Fig. 4.8.a shows the surface view of the lens, and Table 4.1) and also in the epithelia of the surface ectoderm of skin (Figs 4.8.b & c, and Table 4.1). Tissue incubated in the absence of I-1 showed no positive staining (Fig. 4.8.d). This tissue also had a unilaminar, squamous epithelial phenotype (Williams *et al.* 1995).

4.2.10. Confirmation of the inhibitor-1 expression pattern in the kidney

The anti-I-1 antibody from Serotec was deleted from their catalogue after completion of the expression analysis in the mouse whole-mount tissues. Fortunately, an alternative source of I-1 antibody was obtained from Santa Cruz Biotechnologies. This goat anti-I-1 antibody was used to label whole-mount kidneys (following the same protocol for the Serotec antibody) and a similar, though weaker, staining pattern was obtained. This confirmed that I-1 was expressed by the mesothelium in embryonic kidneys.

Cryostat sections of frozen mouse kidney were analysed using the Santa Cruz antibody, but no adequate staining pattern was obtained. It is possible that the I-1 antigen may be lost during the processing of the tissue. This showed the superior tissue resolution obtained from the use of whole-mount tissue and confocal analysis.

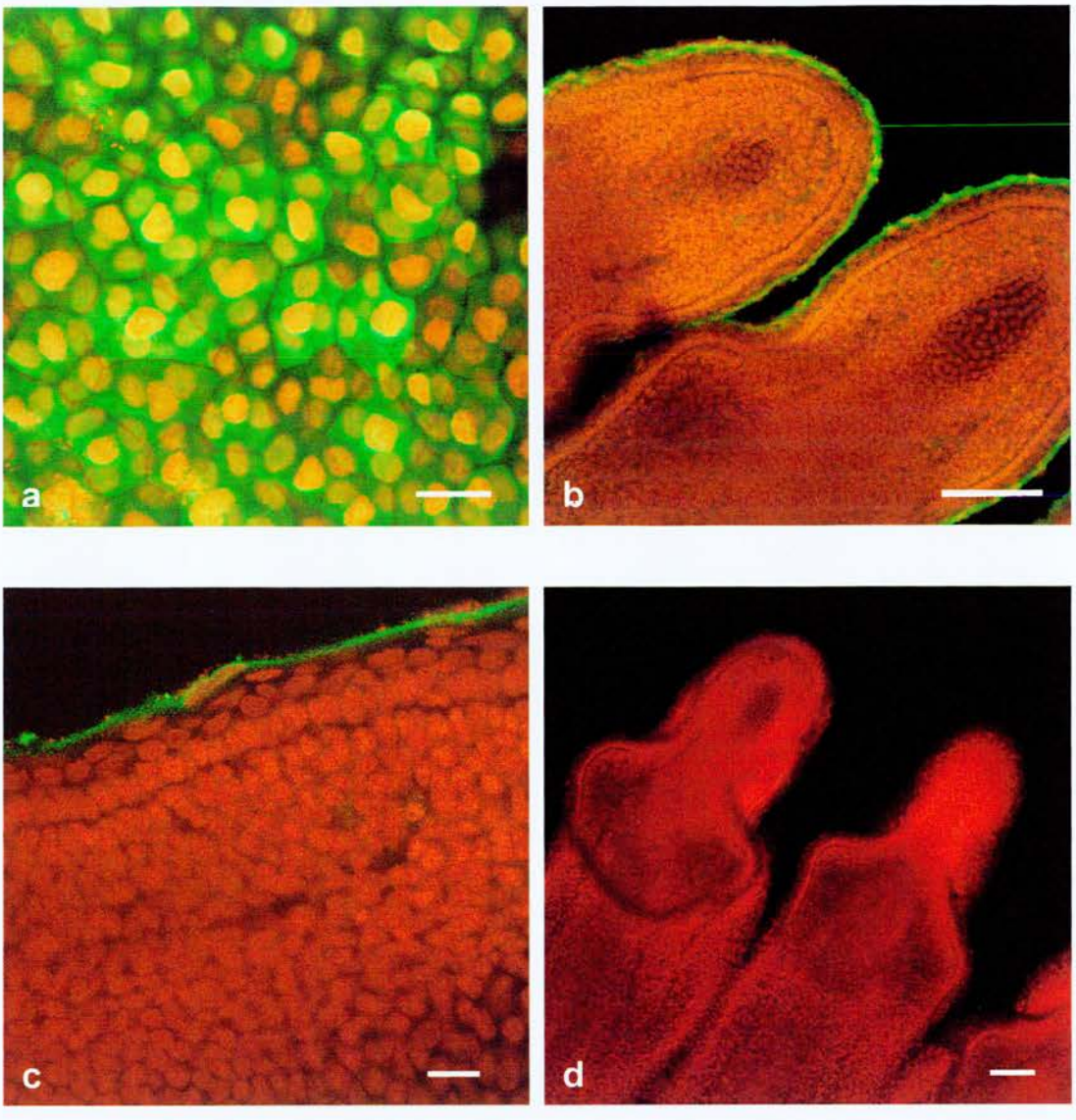


Figure 4.8. Inhibitor-1 expression in whole-mount E16·5 mouse tissue. Inhibitor-1 was visualized with FITC and the tissue was counter-stained with propidium iodide. (a) I-1 expression was present in the surface ectoderm layer of epithelial cells in the eye (surface view) and (b & c) the skin. (d) Incubation of tissue without the I-1 antibody showed no positive staining. Bars: (a & c) 20 μm; and (b & d) 100 μm.

4.2.11. Expression of inhibitor-1 in human lung tissue

The possibility that I-1 could be used as a marker for mesothelium in human tissue and perhaps, as a marker for mesothelioma was investigated using whole-mount tissue and confocal microscopy. A sample of normal whole-mount human lung that had been formalin fixed (supplied by Dr Donald Salter, Department of Pathology, University of Edinburgh) was incubated with the I-1 antibody, but the tissue had very high auto-fluorescence in all channels of the confocal microscope. Analysis of human tissue that had not been incubated in antibody or blocker confirmed that the tissue did auto-fluoresce. This may have been because of the formalin fixative used or, perhaps, it is a feature of the adult lung. Analysis of the unfixed, unprocessed tissue confirmed that the adult tissue had endogenous autofluorescence. This meant that fluorescent confocal analysis was not possible for the human lung tissue. Autofluorescence is a major problem in analysis of material because of the flavoproteins and other cellular components interfering with the fluorochrome in use (Aubin 1979; Benson *et al.* 1979). Andersson *et al.* (1998) used mammalian cells and a confocal laser scanning microscope to show that mitochondria, and probably lysosomes, are responsible for autofluorescence.

Wax embedding of fixed lung tissue followed by sectioning of the material was carried out but I-1 was not detected using an alkaline phosphatase-labelled second antibody. Failure of the tissue to give a signal may reflect a lower level of I-1 expression in human tissue, although this could be a future direction to further investigate I-1 and its expression in human tissue.

4.3. Discussion

The aim of this research chapter was to analyse I-1 expression using confocal microscopy and whole-mount tissue, and this resulted in a more detailed pattern of I-1 expression in comparison to sectioned material.

4.3.1. Inhibitor-1 is a marker for the mesothelium and surface ectoderm

The I-1 expression pattern in the kidney was compatible with the antibody marking mesothelial cells, an epithelial derivative of mesoderm-derived tissue that characterizes the surface of the body cavities (parietal mesothelium) and their constituent tissues (visceral mesothelium). Further analysis revealed I-1 expression in the peripheral epithelia from tissue of the coelom, confirming I-1 specificity to the mesothelium (McLaren *et al.*, 2000, Appendix 2, p. 188). As mentioned in Chapter 1, the lateral plate mesoderm of the developing mouse at E8·5 divides to form the outer derivative (the somatopleure), which lies below the ectoderm, and the inner compartment (the splanchnopleure), which lies next to the endoderm (Williams *et al.* 1995; Kaufman & Bard 1999). Cells lining the coelomic cavity proliferate during development and produce either mesenchymal or epithelial cells: these mesenchymal cells differentiate to form the epithelial lining i.e. the mesothelium (Kaufman & Bard 1999). Other structures formed from these mesenchymal cells include smooth muscle and connective tissue coats of the gut. Localized mesenchymal populations form parts of the liver, spleen and early kidney. Regions of the coelomic epithelium produce the mesonephric epithelium, cells which envelope the primordial germ cells, and the lining of the genital tracts (Williams *et al.* 1995; Kaufman & Bard 1999). Since the results in this chapter report I-1 expression in the coelomic epithelium or mesothelium only (and not its derivatives), this suggests that I-1 may be down-regulated as the tissue differentiates.

The Pax-2 antibody labelled the condensing mesenchyme in the developing kidney and did not co-express with I-1: the anti-I-1 antibody labelled peripheral epithelial cells. Therefore, I-1 is not a stem cell marker in the developing kidney.

Confirmation of the mesothelial phenotype of the I-1-positive cells was obtained by laminin and pan cyokeratin labelling of the kidney and other tissues. The pan cyokeratin expression pattern in the kidney coincided with the I-1-positive cells, and the laminin antibody detected the basal laminae secreted by these cells. However, both antibodies showed weak expression for early kidneys and the

laminin expression pattern did increase as the kidney matured. The reasons for this weak peripheral expression pattern were not entirely clear since the internal epithelia of the kidney strongly labelled with both antibodies. Therefore, the epithelial layer of the kidney weakly expresses laminin and cytokeratin in the early stages of kidney development unlike other organs from the coelomic cavity (e.g. lung). This may reflect a difference in kidney development from other visceral organs, since the kidney arises from the intermediate mesoderm whereas the heart, for example, forms from coelomic epithelial cells (Kaufman & Bard 1999). Although the origin of the kidney's mesothelial layer and capsule is unknown, it is possible that this layer derives from the mesothelial cells lining the coelomic cavity, or it differentiates from the peripheral cells of the metanephric mesenchyme.

There were several reasons to support the hypothesis that I-1 was specific to the epithelium of surface ectoderm and the mesothelium. Incubation of tissue in the presence of serum but in the absence of the I-1 antibody gave no background staining, showing that the expression pattern was not a result of the secondary antibody. Removal of the coelomic epithelium from E11 kidney rudiment showed no I-1 staining of the early kidney, and thus, the I-1 staining pattern was not caused by an 'edge effect'. There was internal I-1 labelling in a small number of testes at ~E13, and thus, lack of penetration of the antibody did not cause the distinctive I-1 staining pattern. Use of an anti-I-1 antibody from a different supplier (Santa Cruz Biotechnology Inc.) showed an identical staining pattern to the Serotec antibody. Furthermore, analysis of kidney sections after wax embedding confirmed that I-1 was present only in the peripheral cells, although the resolution of the tissue in these sections was poor and there was non-specific background staining.

Previous analyses of the I-1 expression have mainly used immunoblotting (MacDougall *et al.* 1989; Hemmings *et al.* 1992; Fryckstedt *et al.* 1993; Sakagami *et al.* 1994; Lowenstein *et al.* 1995). The only published data for the I-1 protein expression in kidney was by Fryckstedt *et al.* (1993), who analysed newborn rat kidney cryostat sections and showed I-1 protein expression in some tubule cells of

the inner cortex. However, these results are in conflict with the work in this chapter and those by Svennilson *et al.* (1995). Wax-embedded mouse kidney sections which had been incubated with secondary FITC antibody only (i.e. no primary antibody was applied; Figs 4.2.d–f) showed non-specific labelling of some cells, similar to those shown in the study by Fryckstedt. The *in situ* hybridization results of E15 rat kidney sections by Svennilson *et al.* (1995) do not correspond to those obtained in this chapter but there was no mention of the use of a negative control (i.e. sense oligonucleotide) used in the above study. Therefore, it may be possible that the I-1 mRNA is synthesized by the peripheral stem cells but expressed by the mesothelial cells.

There was little cross-reaction with other proteins in the whole-mount specimens, as confirmed by low background staining. Also, since the tissues used were derived from a variety of mouse types, the I-1 expression pattern was not strain specific. However, neuronal and skeletal muscle expression were not investigated, but previous research has shown I-1 that is expressed in these tissues (MacDougall *et al.* 1989; Hemmings *et al.* 1992; Sakagami *et al.* 1994; Lowenstein *et al* 1995). The conclusion from this study is that I-1 is not a stem cell marker in the developing kidney.

4.3.2. Inhibitor-1 expression during testis development

During early testes development, cells of the mesonephros migrate into the gonads and form the testes cords (Buehr *et al.* 1993; for review of development, see Kaufman & Bard 1999). This migration been shown to be under the control of the Y chromosome gene *Sry*, which is expressed in mouse from E10·5 to E12·5 (Hacker *et al.* 1995; Martineau *et al.* 1997; Capel *et al.* 1999). The results presented in this chapter using I-1 and laminin antibodies suggested that these migrated cells (or at least a portion of the migrated cells) are epithelial, and from E13·5 onwards, only the surface layer of cells of the testes continued to express I-1. This implied that the I-1 gene may be controlled downstream of *Sry* during testes development, and I-1 expression was lost as the phenotype of these cells altered after migration. Migration of epithelial cells into the testes was also

suggested by Wartenberg *et al.* (1991) after analysis of the developing rabbit testes. Buehr *et al.* (1993) and Martineau *et al.* (1997) showed that several cell populations migrate from the mesonephros into the developing testes. In the developing ovary, there is limited invasion of mesonephric cells in the cortex of the gonad.

4.3.3. Future work

The possibility that I-1 could be used as a marker for mesothelioma in humans has not been fully investigated in this project because of the problem of endogenous autofluorescence of human lung tissue when analysed by confocal microscopy. This is an aspect of the research that could be analysed further. It raises the possibility that I-1 could be used as a specific marker for mesothelioma, a fatal malignant cancer of the mesothelium. The tumour is most commonly found in the pleural cavity since the disease is predominantly linked to asbestos exposure. Diffuse malignant mesothelioma is an aggressive tumour characterized by serosal thickening. The tumours range from epithelial through to mesenchymal in tissue type, and thus, are categorized as being epithelial, sarcomatous / fibrous or biphasic (for review, see Attanoos & Gibbs 1997). As yet, there is not a marker available that can effectively distinguish between mesothelioma and other tumours of the pleura (e.g. adenocarcinoma). Currently, a panel of antibodies is required to verify tumour type, including antibodies against glandular markers (e.g. CEA and B72·3), all of which are negative for mesotheliomas (for review, see Whitaker 2000). Anti-cytokeratin CK 5/6 has been shown to be potentially useful to differentiate the epithelioid tissue of mesothelioma (positive staining) and the adenocarcinoma (negatively stained tissue; Clover *et al.* 1997). Oates & Edwards (2000) found the antibodies calretinin and HBME-1 to be the most successful markers for mesothelioma distinction, although this view differs from other researchers since delineation between tumour types is not absolute. Oates & Edwards (2000) also showed that the gene product of transcription factor WT-1 was detected in 72% of mesotheliomas and 20% of adenocarcinomas, but the protein is not detectable in autopsy material, which is a drawback. Whitaker (2000) suggested antibodies against EMA, Calretinin and Cytokeratin 5/6, and an

additional two antibodies (selected from the group CEA, B72·3, Ber-EP4 and LeuM1) provide a suitable range of markers for mesothelioma distinction.

Clearly, there is a need for a specific marker for mesothelioma identification since the exclusion method currently used is fallible. The epithelial form of malignant mesothelioma is the most common form of the tumour; thus, I-1 may have the potential to be a mesothelioma-specific marker for these forms of the tumour. However, given that the sex cords of the developing testis were only transiently labelled (presumably differentiation occurred after the cord formation), I-1 expression may be down-regulated after mesothelial tissue differentiation.

Chapter 5
Analysis of the 5' flanking region of the mouse inhibitor-1 gene

5.1. Introduction

5.1.1. Abstract

A preliminary investigation of potential regulatory elements of the I-1 gene was undertaken. Varying lengths of 5' flanking genomic sequence were inserted into a reporter gene construct, and introduced into a human embryonic kidney and a rat mesothelial cell line. The results show that the rat mesothelial cell line 4/4 RM-4 had a lower transfection rate than the human embryonic kidney cell line HEK293. No conclusions could be drawn from the transfection experiments since no cells expressed the fluorescent protein. The only difference that may be relevant was that the HEK293 cells gave a higher number of weakly fluorescent cells after transfection with the 3·3-kb genomic construct. Although inconclusive, the groundwork for future analysis has been carried out.

5.1.2. Aim of the chapter

The experiments described in this chapter were designed to determine whether there were transcriptional regulatory elements in the 5' flanking sequence of the mouse I-1 gene. This would either reflect the presence of promoter and/or enhancer elements necessary for protein expression. The initial work carried out involved insertion of 5' flanking sequence into a vector that contained the β-galactosidase, with the aim of producing transgenic mice. Thus, the role of I-1 during development of the whole embryo could then be determined.

5.1.3. Background and strategy for analysis of inhibitor-1 flanking region

5.1.3.1 The pGT1·8IRESβgeo vector: Chapter 3 detailed the screening of a mouse genomic library and the isolation of two types of clones which contained homology to rat I-1 mRNA. These clones were characterized and showed that the type 2 clone contained the first coding exon for the deduced I-1 protein and ~9 kb of flanking 5' sequence that could potentially contain promoter sequences. Therefore, it was planned that the reporter DNA GT1·8IRESβgeo (constructed by

W. Skarnes, P. Mountford and P. Tate in 1992; Fig. A1e, Appendix 1, p. 170) would be ligated to the flanking 5' DNA. This construct would then be introduced into the mouse germ line, resulting in the production of transgenic mice and I-1 expression during kidney development could be followed *in vivo*. This expression pattern could then be compared to I-1 distribution *in vitro*, as determined in Chapter 4.

The pGT1·8IRESβgeo vector was selected for analysis since it contained the βgeo gene and internal ribosome entry site (IRES) sequences. βgeo is a fusion of the *lacZ* gene and the *neo* neomycin resistance gene. The *lacZ* gene allows β-galactosidase detection of the tagged I-1 gene when it is introduced into the mouse germ line. G418 sulphate is an aminoglycoside that disrupts protein synthesis, and is used for the selection of prokaryotic and eukaryotic cells which contain the *neo* gene (Calbiochem data sheet; Darmstadt, Germany). The presence of βgeo in a construct allows both the reporting and selection of a gene when introduced into embryonic stem cells and grown on selective medium. However, gene trapping can be unpredictable but inclusion of the IRES sequences results in the reporter gene being independently translated. IRES are DNA sequences which interact with ribosomes and contribute to efficient RNA translation. This is an advantage since older vectors which lack IRES result in the fusion protein only being generated from the correct frame, and orientation of both the reporter and gene of interest (Chowdhury *et al.* 1997) whereas any insertion into an active gene should give functional reporter expression using IRES. Furthermore, IRES does not have any detrimental effects on gene expression, it is not toxic to the ES cells (Mountford & Smith 1995), it integrates in the genome of the ES cells in an unbiased manner and it is not tissue specific (Chowdhury *et al.* 1997).

Figure 5.1 shows how the genomic clone type 2 (isolated from the library screening in chapter 3) was modified to include *Sal*I restriction sites for the insertion of the GT1·8IRESβgeo DNA and *Xho*I restriction sites for isolation of the final construct from the vector backbone. Therefore, oligonucleotides were designed to contain the restriction sites *Xho*I and *Sal*I.

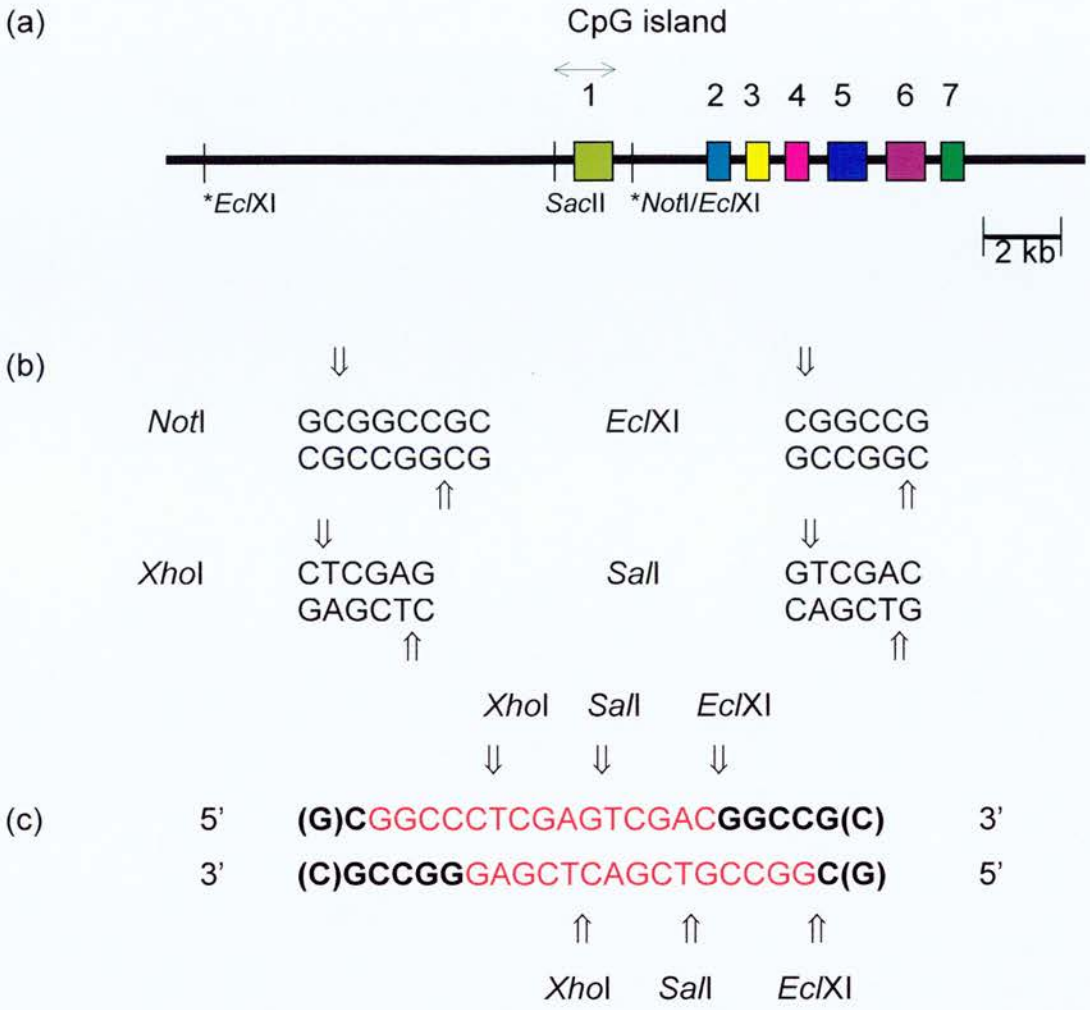


Figure 5.1. Insertion of unique restriction sites into clone type 2 DNA. (a) The type 2 DNA (black line) showing the position of the I-1 protein-coding exons (coloured boxes, not to scale) and the position of the CpG island (not to scale). The restriction sites for insertion of oligonucleotides are indicated by an asterisk (*). The scale bar is indicated. (b) The sequence of the restriction sites of *Not*I, *Ecl*XI, *Sal*I and X*ho*I. The arrows indicate the site of digestion. (c) The oligonucleotide sequences (red) were inserted at the *Not*I and *Ecl*XI sites. The 5' and 3' nucleotides of the *Not*I restriction site are in brackets. Thus, the unique restriction sites of X*ho*I and *Sal*I were introduced into the clone resulting in loss of the *Not*I site but the *Ecl*XI site remained intact.

Digestion of the genomic DNA with *Not*I and *Ecl*XI followed by insertion of the oligonucleotides would result in the presence of the *Xho*I and *Sal*I unique sites within the genomic DNA (Figs 5.2 & 5.3, and Fig. A1e, Appendix 1, p. 170).

However, building of the pGT1·8IRESβgeo construct was not successful since the reporter and genomic DNA did not ligate, even after extensive modifications to both types of DNA and alteration of the experimental protocol.

5.1.3.2. The pEGFP-1 vector: Because of the failure to ligate the reporter GT1·8IRESβgeo and genomic DNA, a different approach for determining the presence of promoter elements was required. As an alternative to GT1·8IRESβgeo, an enhanced green fluorescent protein (EGFP) vector was used. The vector pEGFP-1 (Clontech Laboratories Inc.; GenBank Accession number U55761; Fig. A1f, Appendix 1, p. 171) is promoterless and expresses EGFP only if a functional promoter is present in the inserted DNA. Therefore, insertion of the mouse genomic DNA into this vector followed by transfection of mammalian cells would indicate whether the genomic DNA contained any promoter elements. The advantage of using this vector was that transfected cells could be directly visualized using confocal microscopy, whereas β-galactosidase would be detected indirectly.

GFP has widely been used for direct visualization of proteins in living cells since its isolation from the jellyfish, *Aequorea victoria*, and has been used in the analysis of gene expression and cell lineage (Chalfie *et al.* 1994). There have been various modifications to the wild-type molecule to maximize the emission of the protein produced and to allow its use in mammalian cells grown at 37 °C. The pEGFP-1 vector from Clontech encodes the GFPmut1 variant that has a double amino acid substitution of Phe^{64} to Leu and Ser^{65} to Thr, resulting in a brighter fluorescence than GFP.

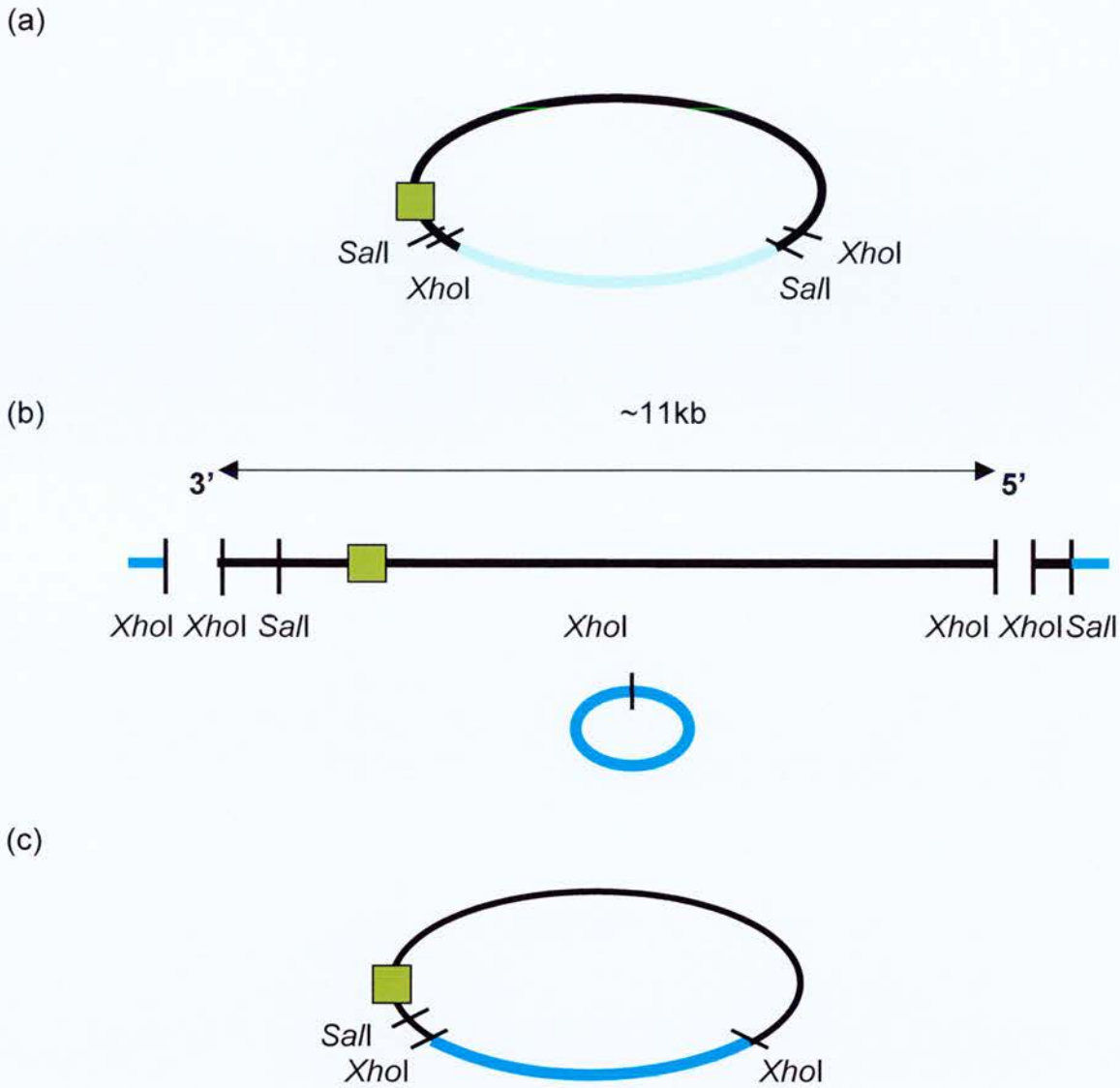


Figure 5.2. The plan for transfer of genomic type 2 DNA into a *Sal*I-knockout vector. Type 2 clone DNA (black line) from the original pBluescript vector (pale blue) was transferred into a *Sal*I-knockout pBluescriptIIKS vector (darker blue). The first exon is represented by the olive box (not to scale). (a) The modified type 2 clone showing the position of the inserted *Xho*I and *Sal*I oligonucleotides. (b) The genomic DNA after digestion with *Xho*I (~11 kb). The *Sal*I knockout pBluescriptIIKS vector was digested with *Xho*I. (c) The final pBluescriptIIKS *Sal*I knockout vector after insertion of the *Xho*I genomic DNA fragment. Only one *Sal*I site is present in the final plasmid.

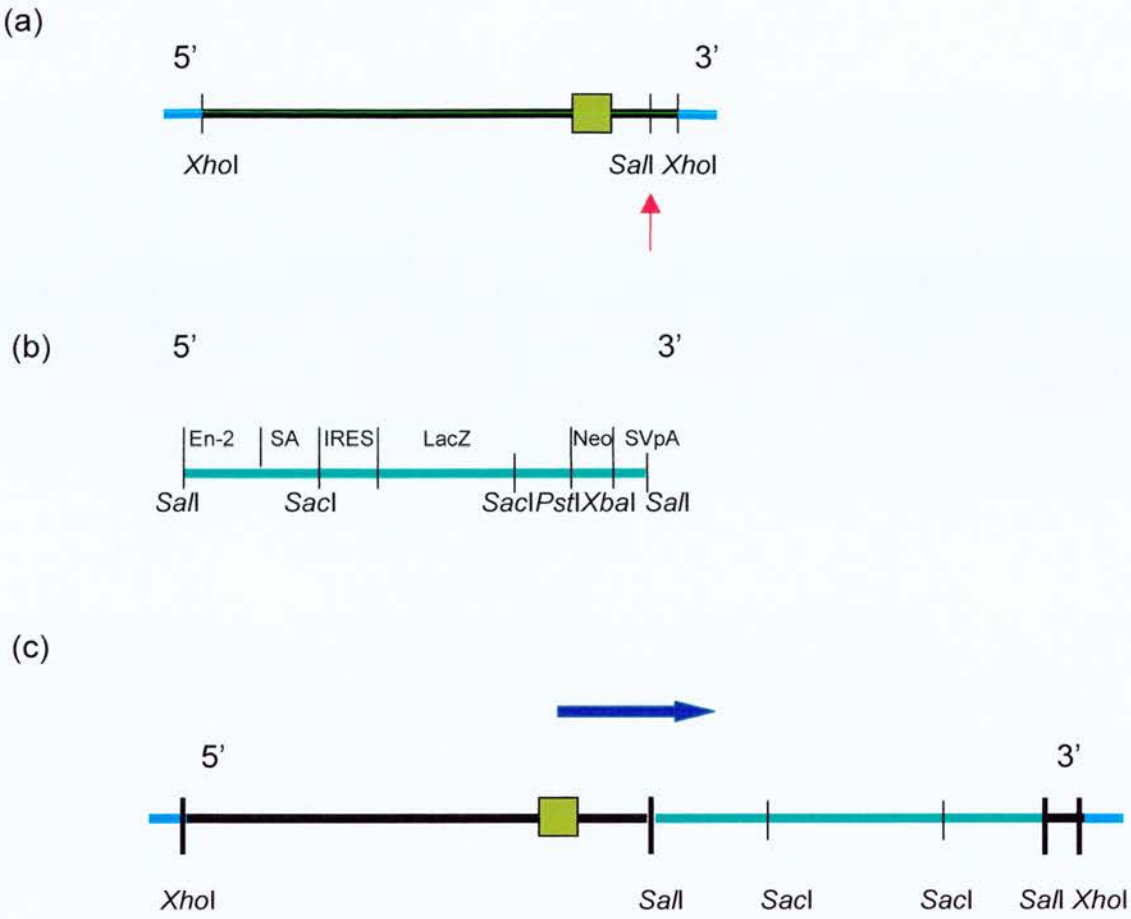


Figure 5.3. The steps involved in building the final GT1·8IRESβgeo construct. The construct was designed to contain the modified genomic DNA (black line; ~11 kb) and the GT1·8IRESβgeo DNA (green line; ~9kb). (a) The linearized pBluescriptIIKS DNA is represented by a blue line and the first exon is represented by an olive box (not to scale). The red arrow shows the position of the unique *Sal*I site into which the GT1·8IRESβgeo DNA would be inserted. (b) The linearized GT1·8IRESβgeo DNA showing the components of the vector and selected restriction sites. (c) The predicted final construct expected after ligation of the insert into the genomic vector.

The altered type 2 mouse genomic clone DNA outlined in 5.1.3.1, above (i.e. after insertion of the oligonucleotides containing the *Xho*I and *Sal*I unique restriction sites), was digested then successfully inserted into the pEGFP-1 vector without further modification. Two constructs containing 9 kb and 3·3 kb of genomic DNA were produced (pEGFPLM9 and pEGFPLM3·3, respectively; Fig. 5.4). These final constructs were then used to transfect cell lines.

5.1.3.3. Transfection of mammalian cells: Transfection is the introduction of foreign DNA into cells to produce a protein and it has become a powerful tool for analysis of gene regulation and function. Transiently transfected cells (or transients) are cells which grow in the absence of a selective agent (e.g. G418 sulphate), but not all the cells will be transfected. Stable-transfected cells survive in the presence of a selective chemical since each cell has been transfected and must contain the resistance gene from the introduced DNA in order to survive.

The method chosen for analysis of the final EGFP genomic constructs was the transfection of mammalian cell lines with the reagent FuGENE 6 reagent (Roche) since the reagent required minimal optimization and is non-toxic to the cells. In this research chapter, transients were analysed using the cell lines HEK293 (derived from human embryonic kidney) and 4/4 RM-4 (a rat mesothelial cell line). HEK293 cells are highly transfectable and have been used extensively in transfection studies, and thus, were the control cell line to show that the experimental system was working optimally. The reason for using the 4/4 RM-4 cell line was that Chapter 4 showed that the I-1 protein was expressed in the mesothelium. Therefore, transfection of the mesothelial cells with the constructs pEGFPLM9 and pEGFPLM3·3 would confirm the presence of I-1 promoter elements.

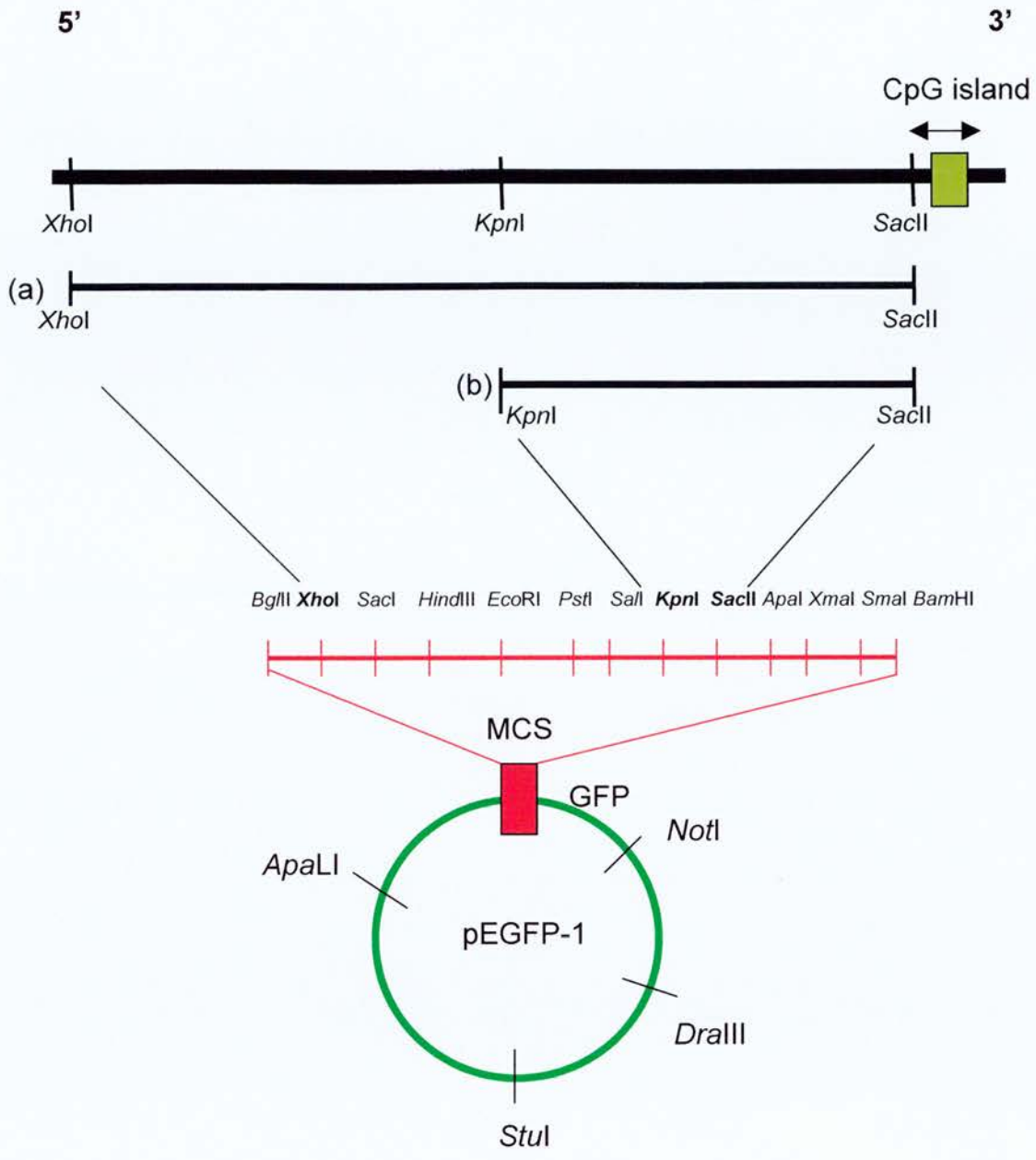


Figure 5.4. The EGFP constructs developed for analysis of the 5' flanking DNA of the inhibitor-1 gene from type 2 clone. The *Sac*II restriction site lies ~300 bp before the start of the first exon (olive box) and just within the 5' end of the CpG island (indicated by the arrow). The exon and CpG island are not to scale. The (a) 9-kb and (b) 3·3-kb lengths of flanking genomic DNA were isolated using enzymes *Xho*I and *Sac*II, and *Kpn*I and *Sac*II, respectively. The restriction fragments were cloned into the multiple cloning site (MCS) of the pEGFP-1 vector.

The vector pEGFP-N1 (Clontech; GenBank Accession number U55762; Fig. A1g, Appendix 1, p. 172) was used as a transfection marker to ensure that the experimental design was working properly. This vector contains a strong promoter (derived from the cytomegalovirus), and thus, allowed each transfected cell to express EGFP. The vector also contains the same GFPmut1 variant as pEGFP-1 for enhanced fluorescence. For negative controls, untreated cells, the addition of FuGENE 6 in the culture medium and transfection with the vector pEGFP-1 were used. All the DNA used in the transfection studies were derived from maxipreps, and therefore, were ultra-pure and contained low levels of endotoxins which could potentially affect cell growth (FuGENE 6 handbook, Roche).

5.2. Results

5.2.1. The GT1·8IRESβgeo construct

The initial work in this chapter involved the modification of the genomic type 2 clone isolated from the library screening in Chapter 3. This involved the introduction of the unique restriction sites X*ho*I and *Sal*I into the DNA at the existing *Not*I and *Ecl*XI sites (Figs 5.1, 5.2 and A6, Appendix 1, p. 180). The *Sal*I site would allow ligation of the GT1·8IRESβgeo DNA into the vector containing the genomic DNA (Fig. 5.3). The unique X*ho*I site would then allow complete isolation of the final construct from the pBluescript backbone before introduction into the mouse germ cells.

Initially, the type 2 clone DNA was digested with *Not*I or *Ecl*XI, followed by ligation of the *Sal*I/*Xho*I oligonucleotides, resulting in the introduction of the X*ho*I and *Sal*I restriction sites flanking the genomic insert of type 2 clone. The *Not*I site was lost after insertion of the oligonucleotides, although the *Ecl*XI site remained intact, as confirmed by restriction digestion. The DNA had to be partially digested with *Ecl*XI to give insertion at the 5' site (Fig. 5.1). Successful insertion of the oligonucleotides was confirmed by restriction digest of the altered type 2 DNA with *Xho*I and *Sal*I since the unmodified DNA did not digest with these enzymes.

This modified genomic DNA was then digested with *Xho*I and the resulting 9-kb fragment was gel purified and GENECLEANed in preparation for the next stage.

After successful insertion of the oligonucleotides into the genomic DNA, it was decided to alter pBluescriptIIKS–/+ to remove the *Sal*I site. There were two reasons for using this vector. First, it was readily available. Secondly, insertion of the 9-kb *Sal*I fragment of GT1·8IRESβgeo would have required partial *Sal*I digestion of the modified genomic DNA and this would have created unnecessary extra steps. Therefore, the most logical plan was to knock out the *Sal*I site from the pBluescriptII vector using *Pfu* polymerase. This enzyme fills in 5' overhangs in the presence of excess dNTPs, and after re-ligation of the pBluescriptII DNA, this resulted in complete removal of the *Sal*I restriction enzyme site in pBluescriptII, as confirmed by restriction digestion. This modified pBluescriptII was then digested with *Xho*I and ligated to the 9-kb *Xho*I genomic fragment (Fig. 5.2). The resulting plasmid DNA contained only one *Sal*I site, into which the reporter DNA could be inserted (Fig. 5.3).

Attempts to ligate the IRESβgeo DNA and the genomic DNA vector were unsuccessful. As an alternative strategy, it was decided to insert the genomic DNA into the partially digested pGT1·8IRESβgeo vector using *Sal*I and insertion of the *Sal*I 9-kb fragment from the modified genomic DNA. However, this method also did not result in the final construct.

Following these efforts, various other modifications were attempted. Insertion of oligonuleotides containing unique restriction sites into both the genomic and pGT1·8IRESβgeo DNA was carried out. This allowed the introduction of unique ends for ligation of the reporter and genomic DNA since the identical *Sal*I ends may have reduced the likelihood of successful ligation. However, no final construct was obtained after this modification. Attempts to insert a unique *Not*I site into the first protein coding exon were unsuccessful using an ExSite PCR-based site-directed mutagenesis kit (Stratagene).

Alternative sources of agarose for DNA extraction, different gel purification methods, transformation of Stratagene Ultracompetent XL-10 Gold cells, use of electro-competent cells and colony screening using the radio-labelled fragment of IRES DNA were used to try to isolate the final construct. Again, these alterations did not result in the ligation of the two DNA types. In each case, the final isolated DNA was either the pBluescriptII/genomic DNA vector or the pGT1·8IRESβgeo vector, depending upon whether the inserted DNA was pGT1·8IRESβgeo or genomic DNA, respectively. It was consistently at the final step where there was failure to ligate the DNA. This avenue of work was exhausted then abandoned since no successful construct was obtained.

5.2.2. The EGFP-1 construct

Due to time and financial constraints, an alternative reporter gene to the IRESβgeo was used. pEGFP-1 is a promoterless GFP vector that has optimized fluorescence and higher expression in mammalian cells than GFP. The multiple cloning site (MCS) of this vector contained restriction sites which were compatible with the sites found in the upstream genomic sequence of the type 2 clone.

The type 2 genomic clone that had been altered to include the unique sites X*ho*I and *Sal*I (as described in section 5.2.1; Figs 5.1 & 5.2) was used to build the EGFP-1 construct. Restriction digestion of this modified type 2 clone with X*ho*I and *Sac*II resulted in ~9 kb of upstream genomic DNA (Fig. 5.4). The 9-kb genomic DNA was digested, gel purified, GENECLEANed then successfully ligated into the X*ho*I and *Sac*II sites of the EGFP-1 vector (Fig. 5.4). The inserted genomic DNA did not include the first exon since presence of the ATG start codon of the I-1 sequence would have interfered with translation of the fusion protein. The unique *Sac*II site at the 3' end of the genomic DNA was ~300 bp upstream from the ATG start codon of the first exon, and the restriction site was within the CpG island (Fig. A7, Appendix 1, p. 181). Ideally, the restriction site used would have been closer to the start of the first exon but lack of restriction sites excluded this possibility: this was confirmed by sequence analysis by the Webcutter

computer program, http://www.medkem.gu.sel. The restriction enzyme *Nco*I recognises the sequence CCATGG and would have been suited to include the entire upstream 5' since it would have recognized the start ATG codon (Fig. A7, Appendix 1, p. 181). Unfortunately, digestion of the type 2 clone DNA with this enzyme showed the presence of many *Nco*I sites in the length of DNA. Also, there was not a corresponding *Nco*I site in the EGFP vector MCS for cloning. Ultimately, there were six identical EGFP clones isolated which had the 9-kb genomic insert present from the 5' *Xho*I site to the 3' *Sac*II site, (pEGFPLM9). These clones were digested and compared to the original type 2 clone DNA to confirm that there had been no loss of DNA during the ligation process.

A second construct (pEGFPLM3·3) was built using *Kpn*I and *Sac*II enzymes to generate a smaller fragment from type 2 clone, and this clone shared the same 3' *Sac*II site with the larger genomic clone (Fig. 5.4). However, the EGFP-1 control used in the ligation reactions (i.e. vector only) that had been digested with *Sac*II and *Kpn*I gave high numbers of background colony numbers, suggesting that the *Sac*II enzyme had not efficiently digested the EGFP-1 DNA. According to the New England Biolabs catalogue (Beverly, MA, USA), certain *Sac*II sites can be resistant to cleavage, and therefore, it is possible that the EGFP-1 *Sac*II site may be resistant to digestion. Therefore, *Ksp*I enzyme, an isoschizomer of *Sac*II, was used to digest the EGFP-1 vector. The 3·3-kb *Kpn*I/*Sac*I fragment was isolated after agarose gel electrophoresis then ligated into the EGFP-1 vector that had been digested with *Kpn*I and *Ksp*I (Figs 5.4 and A6, Appendix 1, p. 180).

Midipreps of the pEGFPLM9 and pEGFPLM3·3 constructs and the vectors pEGFP-1 and pEGFP-N1 were prepared using the Promega Wizard® *Plus* MaxiPrep DNA Purification System. The presence of the inserted DNA was confirmed by restriction digestion and agarose gel electrophoresis.

5.2.3. Transfection using FuGENE 6

FuGENE™ 6 transfection reagent (Roche) is non-toxic, lipid-based and can function in the presence of serum. According to Roche, transfection efficiency is

greater using FuGENE 6 than alternative transfection methods, such as electroporation, and has the added advantage that no major experimental optimization is required. The protocol in the Roche handbook was carried out as described.

Malek & Khaledi (1998) showed that transient transfection of cells using an altered GFP reached maximum expression after 48 h but decreased significantly after 96 h. With this in mind, the initial transfection experiments carried out were cultured for 48 h or 72 h. The positively transfected cells were identified by confocal microscopy after mounting in Vectashield with propidium iodide. Positively transfected cells were labelled green and all cells counter-stained with propidium iodide. The pEGFP-N1 vector was used as a transfection marker since every transfected cell would express EGFP as a result of the presence of the strong CMV promoter. The pEGFP-1 vector was used as a negative control since there would be no fluorescent expression because of the lack of a functional promoter.

The control cell line used was the highly transfectable HEK293 cell line derived from human embryonic kidney cells (a kind gift from Dr Rory Duncan, Department of Biomedical Sciences, University of Edinburgh). These cells were used to show that the transfection conditions were optimal. This cell line has been used extensively for transfection experiments.

The cell line 4/4 RM-4 was obtained from the European Collection of Cell Cultures (ECACC; http://www.ecacc.org.) and is derived from rat mesothelium. There was no mouse mesothelial cell line in the ECACC databank. The rat cell line was used since I-1 is expressed in the mesothelium, as described in Chapter 4. No references to the mesothelial cell line were found in the Roche website (http://biochem.roche.com) or the Medline database (accessed through http://www.bids.com), suggesting that the 4/4 RM-4 cell line had not previously been used for transfection experiments. For each transfection condition analysed, the experiments were repeated two or three times to ensure reproducibility.

5.2.4. Inhibitor-1 protein expression of the HEK293 and 4/4 RM-4 cell lines

To confirm that the I-1 protein was expressed in the mesothelial line, but not the human embryonic kidney cell line, monolayers of the two lines were cultured, fixed in PFA and processed according to the protocol outlined in Chapter 2. The Serotec antibody was deleted from the catalogue shortly after completion of the work carried out in Chapter 4 and the only alternative commercial source of an I-1 antibody was from Santa Cruz Biotechnology.

The results confirmed that the cells received from ECACC were mesothelial (and also reflected the results obtained in Chapter 4). The HEK293 cells did not express the protein, as shown in Fig. 5.5. These results confirmed the suitability of the cell lines selected for the transfection studies.

5.2.5. The plating conditions of the cells for optimal transfection

The transfection protocol for the FuGENE 6 transfection reagent supplied by Roche was followed to determine the best plating conditions for each cell line. The cells were transfected either 6 h or 24 h after trypsinization, and were grown for 48 h. EGFP-N1 DNA was used as a positive marker for transfection. A ratio of 3:1 was used for FuGENE 6:DNA but a ratio of 6:1 was also included for the cells transfected with EGFP-N1. After growth, the cells were fixed in PFA, mounted and viewed using the confocal microscope. The plating condition that gave the better transfection rate was used for the follow-up transfection experiment, where the cells were grown for 72 h after transfection (Table 1).

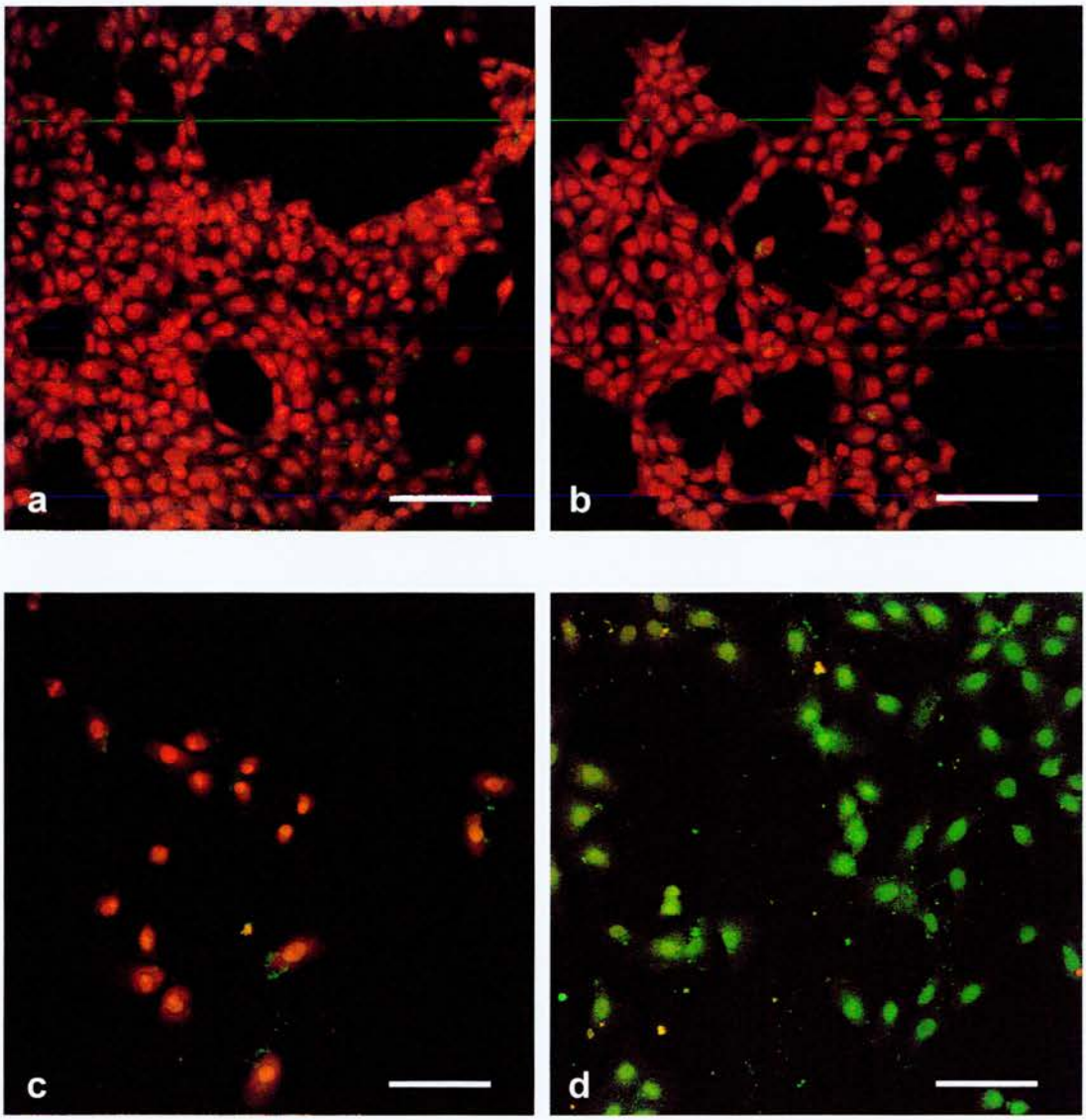


Figure 5.5. Inhibitor-1 protein expression of the cell lines (a & b) HEK293 and (c & d) 4/4 RM-4. The control cells (a & c) were not incubated with the inhibitor-1 antibody. The cells were detected with FITC secondary antibody and counter-stained with propidium iodide (red). The results show that the (b) HEK293 cells do not express inhibitor-1 but (d) 4/4 RM-4 cells do. This confirms the suitability of these two cell lines for use in the transfection experiments. Bars: (a – d) 100 μm.

Table 5.1. Transfection of HEK293 and 4/4 RM-4 cells 6 h or 24 h after trypsinization (indicated by '6 h' or '24 h'). The cells were grown for 48 h or 72 h, fixed in PFA and then analysed by confocal microscopy. The ratio of FuGENE 6:DNA (μL:μg) is shown in brackets. The number of cells was averaged from six random 500 μm^2 areas across the slide.

Treatment	4/4 RM-4 (6 h) grown 48 h	4/4 RM-4 (24 h) grown 48 h	4/4 RM-4 (24 h) grown 72 h	HEK293 (6 h) grown 48 h	HEK293 (24 h) grown 48 h	HEK293 (24 h) grown 72 h
No treatment	0	0	0	0	0	0
FuGENE 6 (3:1)	0	0	0	0	0	0
EGFP-1 (3:1)	0	0	0	0	0	0
EGFP-N1 (3:1)	2·2	2·6	4·5	59·0	60·8	73·3
EGFP-N1 (6:1)	2·8	4·0	6·2	78·4	79·5	93·5

Table 5.1 shows that the plating condition that gave the best number of positive cells for each cell line was transfection 24 h after trypsinization, followed by growth for 72 h. This suggested that there was a difference in fluorescent protein expression in cells cultured longer than 48 h, reflecting the results of Malek & Khaledi (1998).

Fig. 5.6 shows images of the negative controls for both cell lines i.e. cells with no treatment, FuGENE 6 only, and cells transfected with EGFP-1 vector only. There was endogenous fluorescence present in the control cells for both HEK293 and 4/4 RM-4, as shown in Fig. 5.6. These cells provided the basal level to which the highly positive EGFP-N1-transfected cells were compared.

The HEK293 cells transfected with EGFP-N1 gave a high number of positive cells (i.e. cells expressing EGFP) as shown in Table 5.1. Transfection 24 h after trypsinization gave better transfection rates than transfection 6 h after trypsinization. Furthermore, the ratio of 6:1 for FuGENE 6: EGFP-N1 gave higher transfection rate than the 3:1 ratio (Fig. 5.7). The high number of positively transfected cells confirmed that the HEK293 cells are ideally suited for transfection studies.

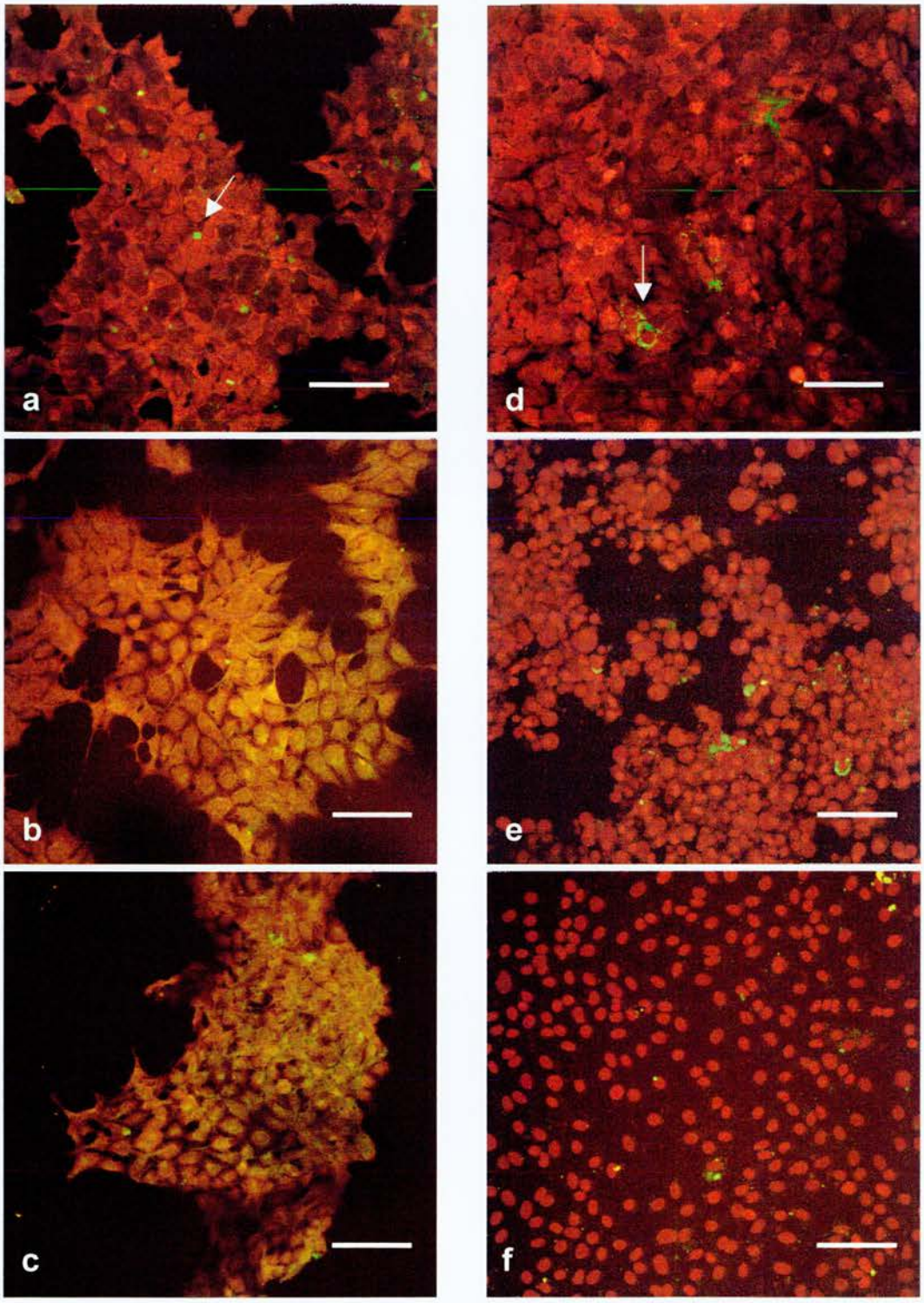


Figure 5.6. The controls used in the transfection experiments for cell lines (a – c) HEK293 and (d – f) 4/4 RM-4. The cells were either: (a & d) untreated; (b & e) grown in the presence of FuGENE 6 or (c & f) transfected with EGFP-1 DNA. The cells were fixed then counter-stained with propidium iodide (red). The white arrows indicate non-specific fluorescent cells. Bars: (a – f) 100 μm.

Table 5.1 shows that the 4/4 RM-4 cells gave far lower numbers of positively transfected cells with EGFP-N1 in comparison to the HEK293 cell line. However, the results also reflect an increase in EGFP-expressing cells when the cells were transfected 24 h after trypsinization using a ratio of 6:1 for FuGENE 6:DNA, similar to the HEK293 results (Fig. 5.7).

5.2.6. The ratio of FuGENE 6:DNA for optimal transfection

The next experiment was designed to vary the ratio of FuGENE 6:pEGFP-N1 DNA to give the optimal number of positively transfected cells. Both cell lines were transfected 24 h after plating and were grown for 72 h, since these were the optimal conditions established in section 5.2.5, above. Table 5.2 shows the effect of the different ratios on positive cell number for HEK293 and 4/4 RM-4 cells, respectively.

Table 5.2. Transfection of HEK293 and 4/4 RM-4 cells using variable ratios of FuGENE 6:pEGFP-N1 DNA ($\mu g\ \mu L^{-1}$). The cells were transfected 24 h after trypsinization, grown for 72 h and fixed before mounting. The ratio of FuGENE 6:DNA (μL:μg) is shown in brackets. The order of transfection rates for HEK293 is arranged from the highest number of transfected cells (1) to the lowest number of transfected cells (5) and was estimated by eye using the × 20 magnification lens. The number of 4/4 RM-4 cells transfected per ratio is indicated and was averaged from six random 500-μm^2 areas across the slide.

Ratio of FuGENE6:DNA ($\mu g\ \mu L^{-1}$)	Order of transfection rate (highest 1; lowest 5) for EGFP-positive HEK293 cells	Average number of EGFP-positive 4/4 RM-4 cells
No treatment	0	0
FuGENE 6 (3:1)	0	0
EGFP-N1 (3:0·5)	5	5·7
EGFP-N1 (3:1)	3	6·8
EGFP-N1 (3:2)	4	6·1
EGFP-N1 (6:1)	1	18·6
EGFP-N1 (6:2)	2	6·6
EGFP-N1 (6:3)	2	3·2

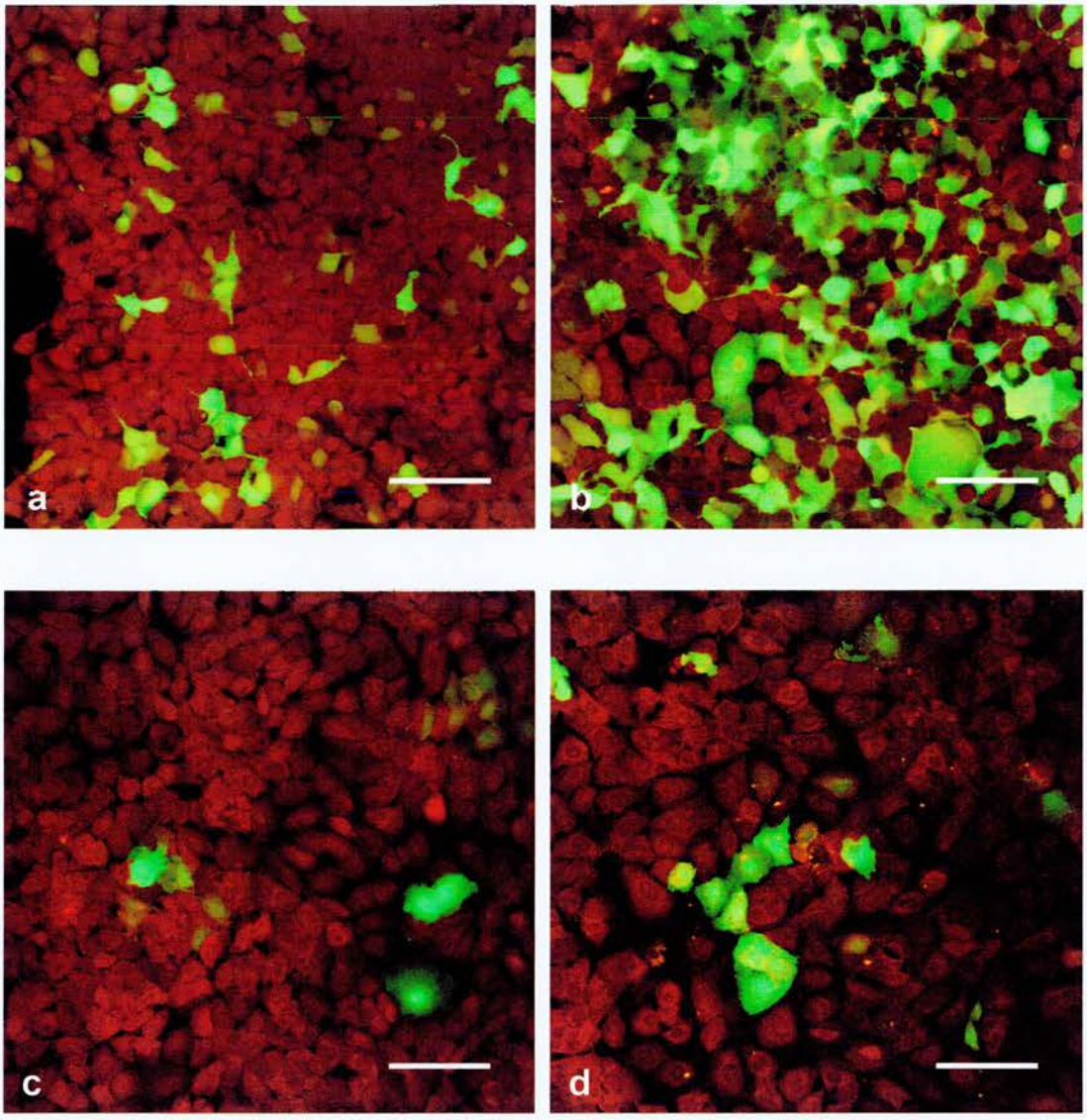


Figure 5.7. Transfection of cell lines (a & b) HEK293 and (c & d) 4/4 RM-4 using EGFP-N1. The positive cells are labelled green (FITC) and the cells were counter-stained with propidium iodide (red). The cells were transfected 24 h after trypsinization, grown for 72 h, fixed then analysed by confocal microscopy. The FuGENE 6:DNA ratios (μg: μL) of (a & c) 3:1 and (b & d) 6:1 were used. Bars: (a – d) 100 μm.

The number of transfected HEK293 cells was very high for all the different ratios of FuGENE 6:pEGFP-N1 DNA and it was difficult to give accurate cell numbers. However, it was possible to compare the different DNA ratios and determine by eye the most effective transfection conditions. Thus, an arbitrary scale with the highest rate of transfection given the values from the highest rate (1) to the lowest rate (6), although even the lowest ranked ratio of FuGENE 6:DNA gave high numbers of positively transfected cells. The results confirmed the preliminary results from section 5.2.5 in that the ratio of 6:1 of FuGENE 6:DNA gave the optimal number of transfected cells for cell line HEK293.

The results for the 4/4 RM-4 cell line showed that the cell line was not as easily transfected as the HEK293 cell line using the transfection marker EGFP-N1. However, the FuGENE 6:pEGFP-N1 DNA ratio of 6:1 gave a significantly higher number of positively transfected cells than other ratios analysed, as shown in Table 5.3, and Figs 5.7 & 5.8. This difference was less pronounced for the HEK293 cells, since each ratio gave high transfection rates.

The experiments in this and the previous section identified the optimal conditions for transfection of the two cell lines. For both the HEK293 and 4/4 RM-4 cells, the highest number of transfected cells were obtained using a FuGENE 6:DNA ratio of 6:1, transfection 24 h after trypsinization followed by growth for 72 h before mounting.

5.2.7. The effect of fixative on fluorescent protein expression

The effect of the PFA fixative on fluorescent protein expression was analysed since it was possible that this could lead to loss of signal. Untreated cells, cells exposed to FuGENE 6, and cells transfected with EGFP-1 and EGFP-N1 DNA, were used to test that there was no overall loss of fluorescent signal.

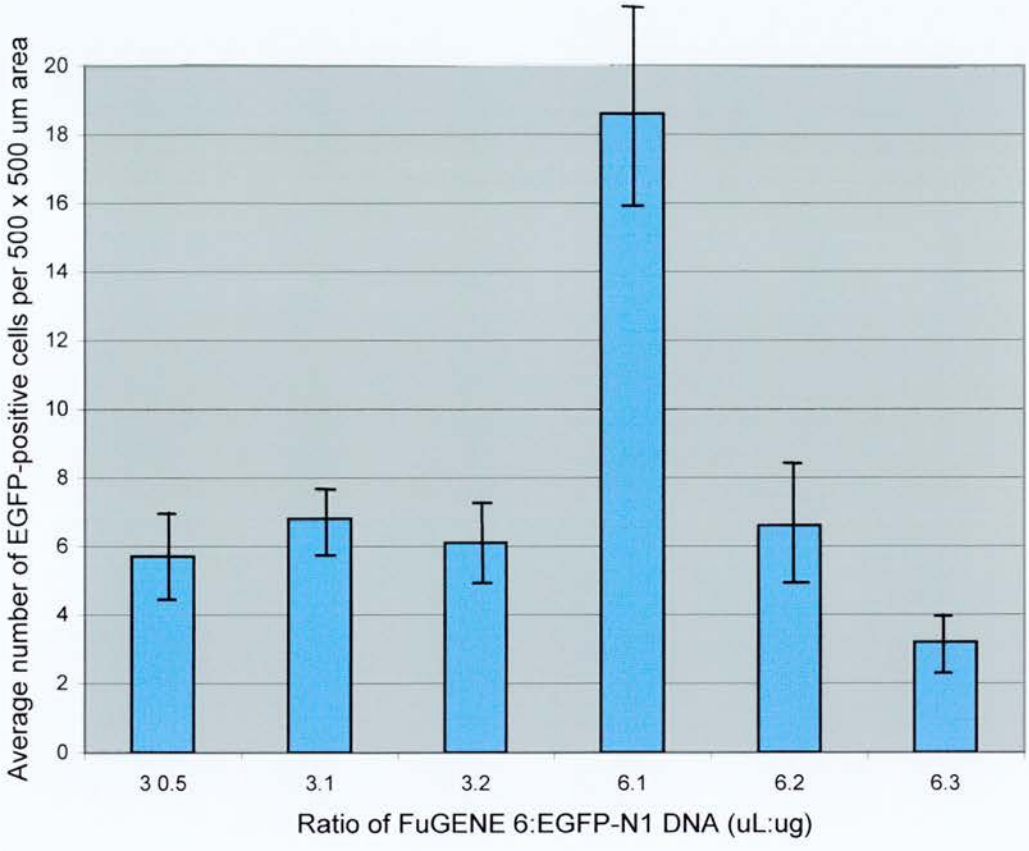


Figure 5.8. The average number of positively transfected 4/4 RM-4 cells using various ratios of FuGENE 6:EGFP-N1 DNA (uL:ug). The error bars indicate standard deviation.

Figures 5.9 and 5.10 show that the fixative had no obvious effect on the number of transfected cells for either HEK293 or 4/4 RM-4 cells, respectively (Figs 5.6 & 5.7). The fixed cells from both cell lines had a dampened signal rather than loss of fluorescence (Figs 5.6, 5.9 & 5.10).

Some of the unfixed cells had a rounded appearance, but there was no loss of signal associated with this. Since the cells were analysed within 2 days of mounting, there was no significant loss of morphology for the unfixed material.

5.2.8. Transfection of cell lines using pEGFPLM9 and pEGFPLM3·3

The genomic constructs pEGFPLM9 and pEGFPLM3·3 were used to transfect the HEK293 and 4/4 RM-4 cell lines using the optimized conditions outlined in sections 5.2.5–7. Both fixed and unfixed cells were analysed since section 5.2.7 showed that the use of PFA appeared to dampen the fluorescent signal, although it did not eliminate fluorescence (Figs 5.9 & 5.10).

Figure 5.11 shows that transfection of both cell lines with the pEGFPLM9 and pEGFPLM3·3 constructs did not result in any positive cells. The positive control EGFP-N1 did give highly positive cells in both cell lines, confirming that the cells lines were transfectable and that the DNA preparation did not cause adverse effect on cell growth. The unfixed negative controls all showed endogenous non-specific fluorescent expression, as shown in Figs 5.9 & 5.10.

The 3·3-kb construct appeared to give a higher number of these weakly positive HEK293 cells, but this was not apparent in the 4/4 RM-4 cell line. This observation may reflect the presence of a weak I-1 promoter. This was the only observable difference between the control transfected cells and those transfected with genomic DNA.

Although the results from the transfection of the cell lines with the genomic constructs were disappointing, the direction for future work is indicated even though lack of time prevented further analysis.

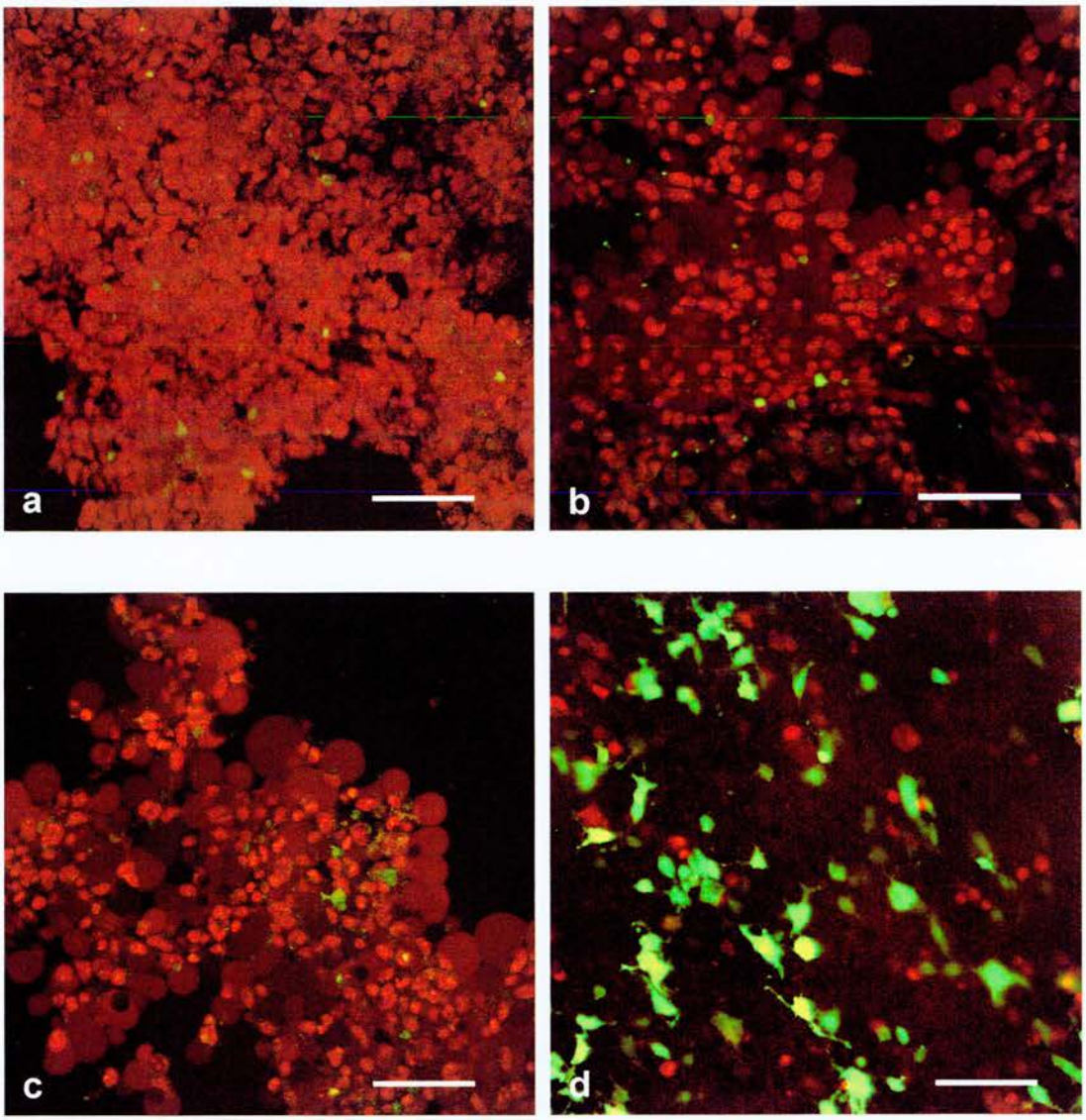


Figure 5.9. The controls used in the transfection experiments for cell line HEK293. The cells were either: (a) untreated; (b) grown in the presence of FuGENE 6; or transfected with (c) EGFP-1 or (d) EGFP-N1. The cells were not fixed before being counter-stained with propidium iodide (red). Bars: (a – d) 100 μm.

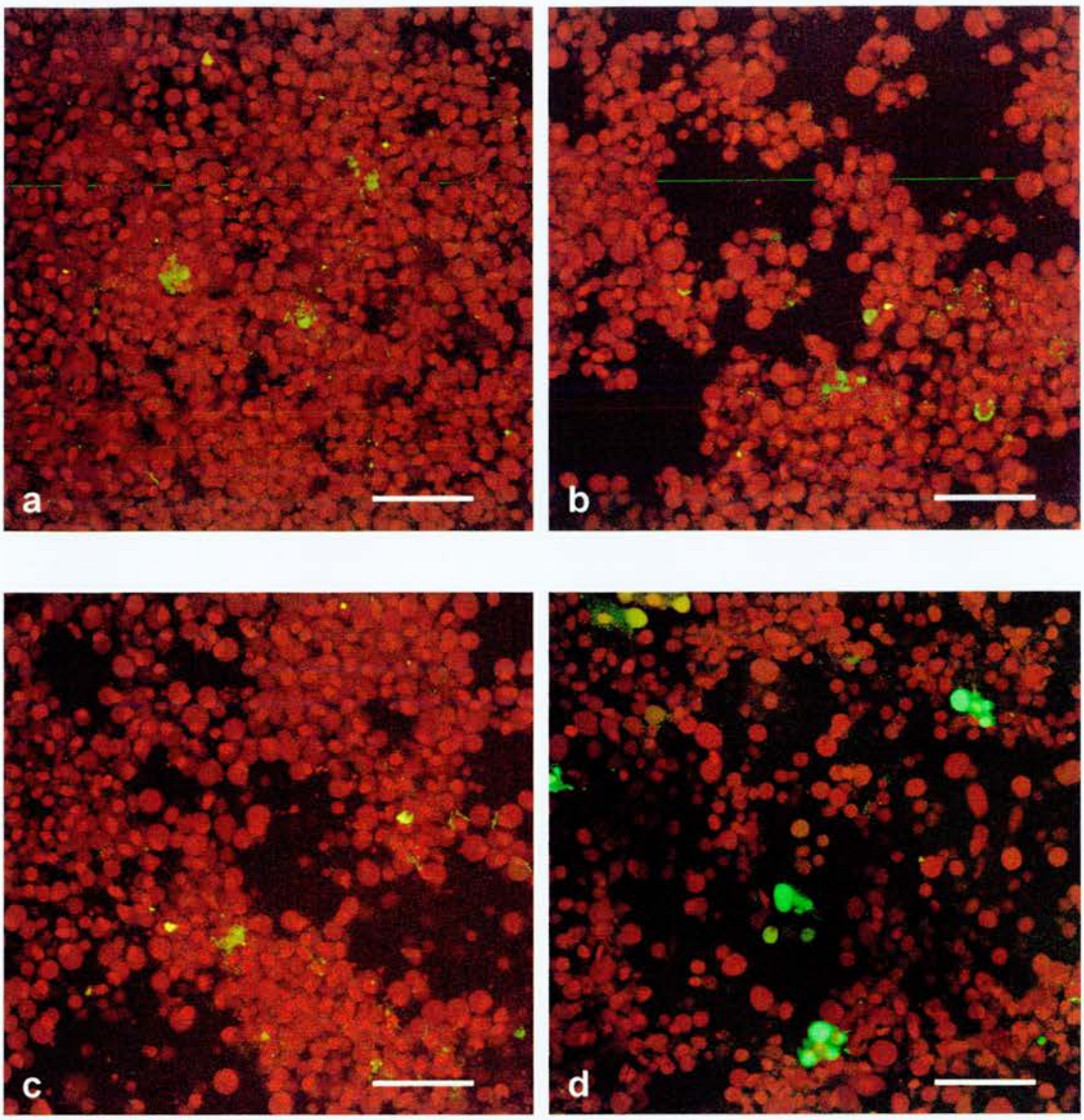


Figure 5.10. The controls used in the transfection experiments for cell line 4/4 RM-4. The cells were either: (a) untreated; (b) grown in the presence of FuGENE 6; or transfected with (c) EGFP-1 or (d) EGFP-N1. The cells were not fixed before being counter-stained with propidium iodide (red). Bars: (a – d) 100 μm.

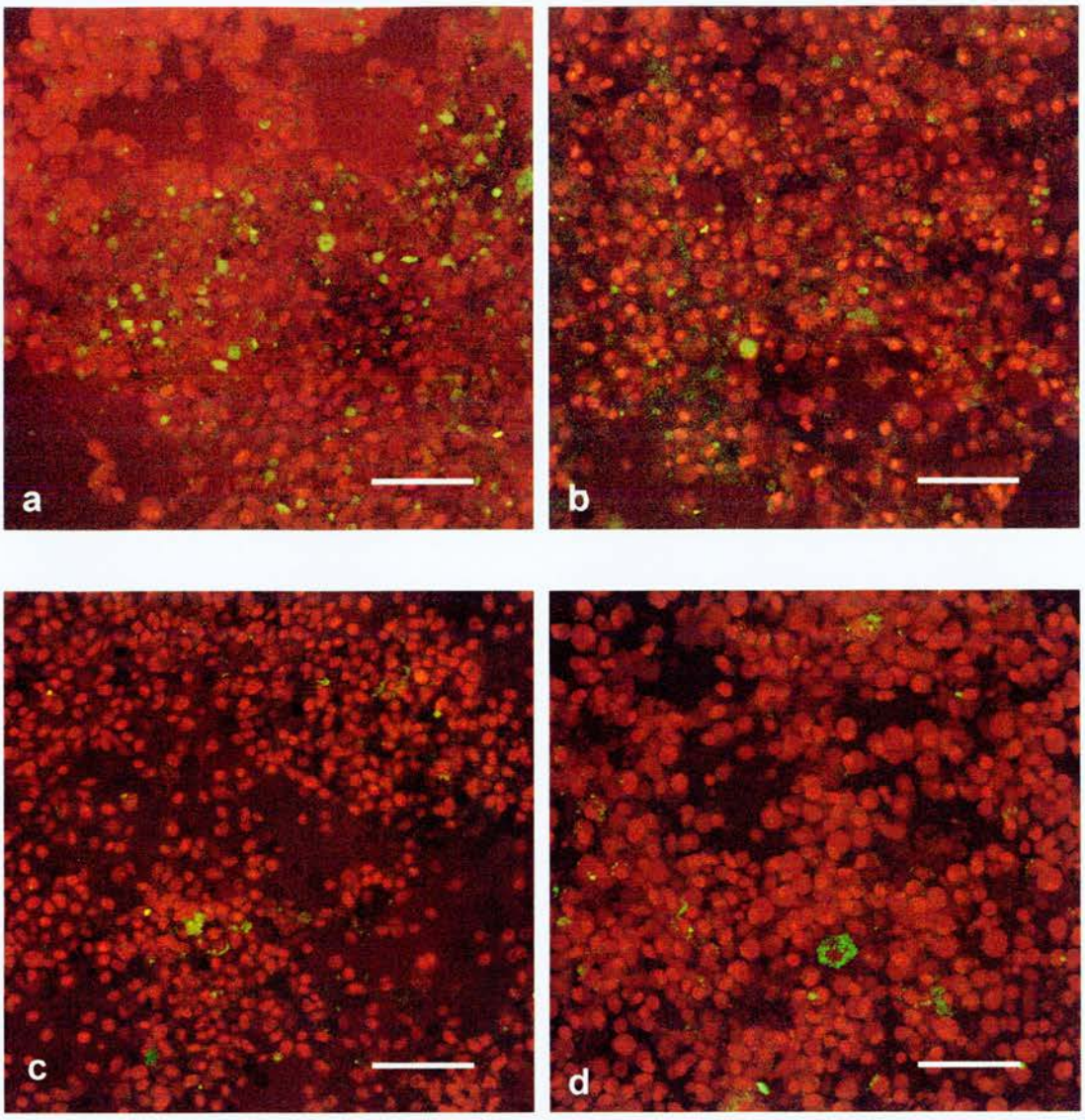


Figure 5.11. Transfection of the cell lines HEK293 and 4/4 RM-4 using the genomic constructs pEGFPLM3·3 and pEGFPLM9. (a & b) HEK293 cells and (c & d) 4/4 RM-4 cells were transfected using the constructs (a & c) pEGFPLM3·3 and (b & d) pEGFPLM9. The cells were not fixed before being counter-stained with propidium iodide (red). Bars: (a – d) 100 μm.

5.3. Discussion

The aim of this research chapter was to characterize the 5' flanking region of the mouse I-1 gene. The building of the construct containing the GT1·8IRESβgeo sequence was not successful, but two lengths of genomic DNA were successfully inserted into a promoterless EGFP vector, resulting in the constructs pEGFPLM9 and pEGFPLM3·3. The initial plan to use the construct to produce transgenic mice was altered because of a lack of time and money, and therefore, transiently transfected cells were analysed. Although inconclusive, the work carried out in this chapter does indicate directions of future work.

5.3.1. Construction of reporter plasmids

The reason for the unsuccessful ligation of the pGT1·8IRESβgeo and genomic DNA is unknown. The genomic DNA was successfully inserted into the pEGFP-1 vector, showing that the genomic DNA was clonable. However, the particular conformation of the pGT1·8IRESβgeo and the genomic DNA may have been unsuitable for successful ligation, or the final size of the construct (>20 kb) was at the upper limit for successful competent cell transformation and growth, and plasmid replication. The size of the predicted final construct may be the most plausible reason. Fortunately, two lengths of the 5' genomic DNA were successfully ligated into the pEGFP-1 vector. These 9-kb and 3·3-kb lengths of genomic DNA were of reasonable size for initial characterization of the upstream sequence of the I-1 gene.

5.3.2. Transfection of cell lines

The cell line 4/4 RM-4 used in this study was confirmed to be mesothelial by its expression of I-1 using the antibody from Santa Cruz Biotechnology. The HEK293 cell line did not express the protein (Fig. 5.5).

In the transfection experiments, there were non-specific fluorescent cells present in the control cells for both HEK293 and 4/4 RM-4 (Fig. 5.6). There was no

difference in the levels of these weakly fluorescent cells in the untreated control group, cells grown in the presence of FuGENE 6 or those transfected with EGFP-1. These control cells provided the basal level to which the highly positive EGFP-N1-transfected cells were compared (Fig. 5.7).

The initial transfection experiments were important to determine the optimal transfection conditions for each cell line. Both the HEK293 and 4/4 RM-4 cells shared optimal conditions: the cells were transfected 24 h after trypsinization using a ratio of 6:1 for the FuGENE 6:DNA, and the cells were grown for 72 h before mounting. PFA fixation of the cells gave a dampened, though not reduced, fluorescent signal. If the I-1 promoter was weak, then any loss in signal may be significant, and thus, both fixed and unfixed cells were analysed when transfected with the genomic constructs. The only minor drawback in not fixing the cells was the development of a rounded shape, but the cells were viewed within 2 days of mounting to prevent major loss of morphology. Each experiment was repeated at least twice to confirm the results presented in this chapter.

In theory, the results from the genomic construct transfection experiments would indicate one of the following:

(1) if both the constructs showed EGFP expression in transfected cells, then the promoter elements will be present in the 3·3-kb length of DNA;
(2) if only the 3·3-kb construct gave GFP expressing cells, then the promoter regions will lie within this region, but there will be negative regulatory sequences upstream (contained within the 9-kb region);
(3) if only the 9-kb construct gave positively transfected cells then the promoter regions must lie upstream from the 3·3-kb DNA sequence; or
(4) if neither construct results in GFP expression then one of the following may be involved:
    (a) the regulatory sequences may lie within, or upstream or downstream from the protein-coding exons;
    (b) there may be negative regulatory elements within the 3·3-kb length of DNA;
    (c) the I-1 promoter is weak;

(d) the 285 bp of 5' upstream sequence between the *Sac*II site and the first exon contains regulatory elements; or

(e) there is tissue specificity of I-1 expression.

Since the positive and negative controls worked well in the transfection experiments using the two cell lines, the lack of positively transfected cells was not a result of the low transfection rates of the cell lines used. The lack of positive cells after transfection with the genomic constructs suggests the reason for this is one of the reasons given in 4 (a–e), above. It would be impossible to determine the precise reason without further experimentation.

The results presented in this chapter suggest that the HEK293 cells transfected with the 3·3-kb construct may have a higher number of the weakly transfected cells. This may not be significant, but does, perhaps, indicate that there may be weak expression in the HEK293 cells using this construct.

The 4/4 RM-4 cells gave lower transfection rates than the HEK293 cells using the EGFP-N1 DNA, reflecting the highly transfectable nature of the HEK293 line, but also suggesting that line 4/4 RM-4 is poorly transfectable. There is no published record for this mesothelial cell line being used in transfection experiments (confirmed by a search of the Roche website and Medline database), perhaps because of low transfectability. Also, the low rates were not a result of the endotoxin levels or purity of the DNA used since the HEK293 cell line gave high transfection rates using the same DNA maxipreps. The basal number of non-specific fluorescent cells was consistent for the controls and genomic constructs, and the 3·3-kb construct did not give rise to any positively transfected 4/4 RM-4 cells. However, this may reflect the poor transfectability of the mesothelial cell line rather than absence of promoter sequences.

The use of the rat mesothelial line could be the reason for the lack of positively transfected cells since mouse genomic DNA was used in the transfection experiments, but there was no mouse mesothelial cell line available from ECACC. The predicted mouse and rat I-1 mRNA and protein sequences are over 95%

homologous (as shown in Chapter 3) although this does not necessarily imply that the regulatory sequences will be strongly homologous.

5.3.3. Sequence analysis

The GRAIL/polII program forms part of the NIX analysis at MRC HGMP-RC (http://menu.hgmp.mrc.ac.uk/menu-bin/Nix/Nix.pl) and was used to examine the 1·8-kb sequence upstream from the first exon. The analysis showed the sequence from nucleotide positions 331 to 365 (Fig. A7, Appendix 1, p. 181) shows the presence of TATA and cap sequences. This analysis gave a figure of 0·779 ('good' according to the analysis) on a scale from 0 (indicating no promoter sequences) to 1 (indicating the likelihood of the presence of promoter sequences). The program recognizes only PolII promoters with TATA-like elements and detects approximately 60% of PolII promoter regions with TATA-like elements.

An alternative algorithm within the NIX program, TSSW, predicts potential transcription start sites using a linear discriminant function (LDF) that combines functional motifs and oligonucleotide composition. LDF values above 4 represent candidate transcription sites. The program detected a 'good' potential functional motif, GA, at nucleotide position 1772 (Fig. A7, Appendix 1, p. 181) with a LDF value of 10·52 that lies within the CpG island and 26 base pairs before the start of the first protein-coding exon. Unfortunately, this particular stretch of DNA was omitted from the EGFP genomic constructs because of a lack of restriction sites present in the genomic DNA. A 'marginal' functional motif was detected at nucleotide position 2335 (Fig. A7, Appendix 1, p. 181) and had a LDF value of 8·81, but this stretch of DNA was not present in either of the genomic constructs.

The results from the NIX computer analysis suggested the presence of potential regulatory elements in the sequenced 1·8-kb length of genomic DNA. However, it was not possible to determine if these features are important for I-1 gene regulation on the basis of the transfection experiments carried out in this research project.

5.3.4. Summary

The purpose of this research chapter was to identify the promoter elements of the I-1 gene. The initial work plan had to be altered as a result of unforeseen circumstances since the building of the construct containing GT1·8IRESβgeo was unsuccessful. The alternative strategy involving the pEGFP-1 vector did result in the production of two constructs. Transiently transfected HEK293 or 4/4 RM-4 cells showed that there were no obviously transfected cells using either the long or the short genomic constructs. Although the results presented in this chapter are not conclusive, neither are they negative since the positive controls determined the transfectability of the cell lines used and do give direction for future research.

5.3.5. Future work

Ideally, the EGFP-1 genomic constructs would have been a result of non-directional cloning since this would have the advantage of insertion of the genomic DNA into the vector in both directions. Lack of unique restriction sites within the type 2 clone (Fig. A6, Appendix 1, p. 180) and time constraints did not allow for further experimentation into this. Thus, future work could involve this line of research.

The 285-bp nucleotides from the *Sac*II restriction site to the start of the first protein coding exon was lost because of restriction site limitation, as confirmed by sequence analysis by the Webcutter computer program. Therefore, it is possible that this region of DNA contains the start site necessary for transcription, as suggested by NIX sequence analysis. Unfortunately, introduction of a unique restriction site into the first exon using ExSite PCR-based site-directed mutagenesis was not successful. Insertion of a unique restriction site nearer to the ATG start codon would have confirmed whether this 285-bp stretch of DNA from the *Sac*II site to the first exon contains (or is part of) the I-1 promoter sequence.

Future work could involve the building of a smaller construct (~1·5 kb) to determine whether there are any negative regulatory elements present in the 3·3-kb length of genomic DNA, but it should include the missing 285-bp sequence. This

would allow the creation of transgenic mice using the EGFP constructs. Furthermore, use of a longer length of mouse genomic DNA (e.g. a YAC clone) would allow the promoter and associated enhancer elements to be identified.

Chapter 6
Final discussion

6.1. Summary

In order to characterize the mouse I-1 gene, the aim of the work undertaken in this project was to:

- isolate genomic clones from a mouse library, from which the genomic structure of the I-1 gene could be determined (Chapter 3);
- analyse the protein expression pattern in the developing mouse embryo using whole-mount tissue and confocal microscopy (Chapter 4);
- characterize the flanking 5' genomic sequence to determine the presence of regulatory elements (Chapter 5).

The rabbit, rat and human protein sequences had previously been determined either by sequence analysis (rabbit) or by prediction from mRNA sequences (rat and human). The mouse sequence had not been determined and, therefore, Chapter 3 described the screening of a mouse genomic library to isolate clones that contained sequence homology to the known rat I-1 mRNA sequences. Analysis of these clones determined the genomic structure of the I-1 gene, including the predicted mouse I-1 mRNA and protein sequences which are over 95% homologous to the rat sequences, and share certain sequence motifs that are required for I-1 binding to PP-1. Chapter 4 showed that the I-1 protein was not expressed in the metanephric mesenchyme and, thus, I-1 is not a kidney stem cell marker: the I-1 protein is a novel marker for the mesothelium and is also expressed in the surface layer of cells of the ectoderm. Chapter 5 reported the analysis of the upstream genomic DNA for potential promoter sequences; the results were inconclusive, but reasons are considered for this. Since each research chapter discussed the results that it contained, this final discussion chapter briefly recaps the findings of each chapter and considers future work based on the work carried out.

6.2. The protein-coding region of the inhibitor-1 gene

Chapter 3 determined that the protein-coding region of the mouse I-1 gene spans a 7 kb region of genomic DNA, with the protein encoded by seven exons. The predicted protein sequence is strongly homologous to the predicted sequence from rat muscle cDNA (Elbrecht *et al.* 1990), the sequenced proteins derived from rabbit skeletal muscle (Aitken *et al.* 1982) and liver (MacDougall *et al.* 1989), and the deduced human sequence from cDNA clones isolated from a brain library (Endo *et al.* 1996). Indeed, the mouse and rat predicted mRNA and protein sequences share over 95% homology.

There are two elements in the human I-1 protein structure which are required for efficient PP-1 inhibition: a KIQF motif at the amino end of the sequence and a threonine at amino acid position 35 (Endo *et al.* 1996). Both features are present in the mouse sequence. There is also a serine at amino acid 67 in the mouse sequence that may be a site of phosphorylation, as suggested by Huang & Paudel (2000) using recombinant human I-1. The apparent molecular mass of mouse I-1 determined by SDS-PAGE is 30·5 kDa (Hemmings *et al.* 1992) and this is consistent with the 30 kDa protein detected for rat I-1 (MacDougall *et al.* 1989; Hemmings *et al.* 1992; Fryckstedt *et al.* 1993; Lowenstein *et al.* 1995) but the ~5% variation in amino acid sequence between mouse and rat mRNA identified in this project would account for the 0·5 kDa difference in protein weight.

6.3. Inhibitor-1 protein expression in the developing mouse embryo

The I-1 protein expression pattern of the mouse embryo was determined in Chapter 4 using whole-mount tissue, a commercially available antibody and confocal microscopy. This is the first time that I-1 expression was determined in this way. The only published data showing I-1 protein detection is by Fryckstedt *et al.* (1993), who analysed cryostat sections of rat newborn kidney. However, that study showed that only the occasional proximal convoluted tubule cell was labelled, and the expression pattern was not consistent with the results in this

project, nor with those of Svennilson *et al.* (1995) who used *in situ* hybridization to show mRNA expression in the stem cells. The results of my detailed project did not show protein expression in these cells, thus, I-1 is not a stem cell marker of the developing metanephric kidney.

The results of this project showed that I-1 was only found in the mesothelium and the surface ectoderm but not in other epithelial cells. Both tissue types have unilaminar epithelial morphology (Williams *et al.* 1995) but neither smooth muscle nor brain tissue was analysed since expression in these tissues has already been documented by MacDougall *et al.* (1989), Hemmings *et al.* (1992), Sakagami *et al.* (1994) and Lowenstein *et al.* (1995). Confirmation of the I-1-positive cells having an epithelial phenotype was shown using epithelial-specific antibodies (anti-laminin and anti-pan cytokeratin). Other tissues of the coelomic cavity, which are bounded by mesothelium, were analysed, showing similar I-1 staining to that found in kidney. This observation suggests a unique expression pattern for I-1 in the embryo and indicates that I-1 could be used as a marker for the mesothelium. Interestingly, the only internal labelling exhibited by I-1 was in the occasional developing testis (~E13) during formation of the primary sex cords, but the I-1 signal was down-regulated after formation of the cords. This suggests that I-1 is downstream from the gene *sry* that controls the migration of mesonephric cells into the developing testis to form the sex cords.

The presence of I-1 in the mesothelial and ectodermal cells will inhibit the activity of PP-1 (I-1 is thought to have no other role in the cell) and so lead to an increase in intra-cellular phosphorylation reactions, although it is not clear if these cells require this inhibition for control of the cell cycle, protein synthesis or some other cellular pathway. The results given in chapter 4 showed that I-1 expression is unaltered as the coelomic epithelium differentiates to mesothelium (determined during kidney development from E11 onwards) though there was loss of the I-1 protein after migration of mesonephric tissue to form the testis sex cords. Since I-1 is expressed in coelomic epithelium and ectoderm from early development through to the adult, these tissues must require I-1 expression for normal function. Moreover, exogenous inhibition of PP-1 (and PP-2A) in cultured E13 rat kidneys

by Svennilson *et al.* (1995) resulted in inhibition of growth and disrupted nephron formation, suggesting that endogenous I-1 and PP-1 levels must be balanced for maintenance of normal growth and development.

Identifying any unique role for I-1 is further complicated by there being other known inhibitors of PP-1. DARPP-32 (dopamine- and cAMP-regulated phosphoprotein) and inhibitor-2 have both shared and separate motifs required for binding PP-1 (Huang *et al.*, 1999). DARPP-32 is expressed in developing and mature rat kidneys (Fryckstedt *et al.* 1993). However, it is not possible to comment on the phosphorylation state of I-1 because the Serotec antibody used in this study can detect both forms of the protein (MacDougall *et al.* 1989). The precise function of I-1 in the developing embryo will be determined by gene targeting experiments.

Allen (*et al.* 2000) analysed wild-type mice and found I-1 expression in the dentate gyrus and hippocampus. The I-1 knock-out mouse (Allen *et al.* 2000) exhibited no major physical defects due to loss of the gene's activity, but there was disruption in long-term potentiation at perforant path-dentate granule cell synapses. Loss of I-1 in the knock-out mice did not effect the insulin levels on glycogen synthase since wild-type mice had identical levels, suggesting that I-1 is not involved in the signal transduction pathway for glycogen synthesis (Scrimgeour *et al.* 1999). This observation suggests that loss of I-1 expression may be compensated for by the presence of another protein phosphatase type-1 inhibitor. However, Scrimgeour found no expression of DARPP-32 in the muscles of the *I-1*$^{-/-}$ mice; thus, there was no increase in levels of this protein after loss of I-1 expression. Also, muscles from mice that lacked both the I-1 and DARPP-32 genes had similar glycogen synthase levels to wild-type mice.

Although MacDougall *et al.* (1989) did not detect I-1 expression in mouse liver by immunoblotting, this was probably due to the lower proportion of mesothelial cells in this tissue in comparison to e.g. kidney and gonad: Fig. 4.6.e clearly shows the presence of I-1 in mouse liver.

Incubation of tissue from the E11 coelomic cavity suggested that the mesothelium lining of the kidney derives from coelomic epithelium (which surrounds the intermediate mesoderm at ~E11·5 when the kidney is still small and envelops the kidney when it enlarges). However, why this should result in a weakly labelled epithelial cell layer was not clear.

6.4. The 5' flanking sequence of the inhibitor-1 gene

The plan of work for chapter 5 (analysis of 5' flanking DNA sequences) had to be altered since the GT1·8IRESβgeo DNA was not successfully ligated to the genomic DNA for reasons unknown, although the final construct size (>20 kb) remains the most likely cause. A different approach was required and this involved the insertion of two lengths of genomic DNA into a promoterless EGFP vector, followed by transient transfection of two cell lines.

The transfection studies using either of the genomic constructs did not result in positive transiently transfected cells in the rat mesothelial cell line 4/4 RM-4 or in the highly transfectable cell line HEK293. The only potentially significant finding was that there was a slight increase in weakly positive cells in the HEK293 cell line after transfection with the 3·3-kb construct. This may not be relevant since it is possible that the genomic construct may cause apoptosis and this could have resulted in this observation. However, the experiments did show that the alternative approach undertaken to analyse promoter sequence analysis was reasonable i.e. use of the promoterless EGFP vector followed by transient transfection of cultured cells. The experiments showed that neither the experimental procedure nor the cell lines used were responsible for the negative results, since the transfection marker EGFP-N1 resulted in positive cells for both cell lines. However, the design of the EGFP constructs was limited by lack of suitable restriction sites in the genomic DNA sequence and attempts to insert a unique site using oligonucleotides and PCR-mediated ExSite mutagenesis were unsuccessful.

The reason for lack of positively transfected cells after transfection with the genomic constructs may have been due to the absence of promoter sequences or the presence of negative regulatory elements in the genomic DNA. The missing 285 bp length of genomic DNA (from the start codon of the first protein-coding exon to the 5' *Sac*II site) may be vital for promoter regulation due to the presence of a transcription start site, as suggested after analysis by the computer program NIX. It is difficult to confirm the functional role of this site without further experimentation. However, the results did show that the experimental procedure was not deficient since positively transfected cells were obtained using a positive transfection marker (EGFP-N1) and the two cell lines. Nonetheless, the experiments described in Chapter 5 do lay the groundwork for future work.

6.5. Conclusion

The aim of the project has been successful in characterizing the mouse I-1 gene: the structure of the gene, and the mRNA and protein sequences have been predicted from genomic clones; the protein expression pattern has been determined by confocal microscopy and whole-mount tissue and has shown a unique expression pattern for I-1 in the developing mouse embryo; and although incomplete, the initial groundwork has been carried out to determine the I-1 gene regulatory elements.

6.6 Future work

There are two areas identified from the work carried out in this research project that can be considered for future work. Firstly, the presence of the I-1 protein in the mesothelium suggests the potential for the protein to be a marker for mesothelioma tissue. However, since there was loss of I-1 expression after migration of the mesonephric cells into the developing testis to form the sex cords, then differentiation of the mesothelium may result in loss of I-1 expression. Future work would confirm this, although use of fluorescent confocal microscopy

may not be the ideal tool since the human tissue has high levels of endogenous fluorescence.

The second area that would benefit from future work is the characterization of the I-1 gene regulatory elements. Although the I-1 protein-coding region of the gene has been fully characterized in this project, the regulatory elements have not been identified. Lack of time and unforeseen problems meant that one of the initial aims of the project (i.e. to produce transgenic mice) was not successful. Inhibitor-1 knockout mice have already been produced by Allen *et al.* (2000) but the mice did not exhibit developmental defects since the mice survived after birth, but this does not invalidate the possibility of producing transgenic mice to show the *in vivo* expression pattern of the I-1 protein. Analysis of the transients after transfection with the genomic constructs, pEGFPLM9 and pEGFPLM3·3, showed lack of EGFP expression in either of the two cell lines. This may have been a result of absent (or incomplete) promoter elements, loss of the transcription site or the presence of negative regulatory elements. Future work could include the use of a YAC clone, for example, that contains a longer length of flanking 5' sequence since the promoter or enhancer elements may be further upstream from the protein-coding sequences.

References

Aitken A., Bilham T. & Cohen P. (1982) Complete primary structure of protein phosphatase inhibitor-1 from rabbit skeletal muscle. *European Journal of Biochemistry* **126,** 235–246.

Allen P.B., Hvalby O., Jensen V., Errington M.L., Ramsay M., Chaudhry F.A., Bliss T.V.P., Storm-Mathisen J., Morris R.G.M., Andersen P. & Greengard P. (2000) Protein phosphatase-1 regulation in the induction of long-term potentiation: Heterogeneous molecular mechanism. *Journal of Neuroscience* **20,** 3537–3543.

Andersson H., Baechi T., Hoechl M. & Richter C. (1998) Autofluorescence of living cells. *Journal of Microscopy* **191,** 1–7.

Aperia A., Bertorello A. & Seri I. (1987) Dopamine causes inhibition of Na^+-K^+-ATPase activity in rat proximal convoluted tubule segments. *American Journal of Physiology* **252,** F39–F45.

Attanoos R.L. & Gibbs A.R. (1997) Pathology of malignant mesothelioma. *Histopathology* **30,** 403–418.

Aubin J.E. (1979) Autofluorescence of viable cultured mammalian cells. *Journal of Histochemistry and Cytochemistry* **27,** 36–43.

Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A. & Struhl K. eds (2000) *Current Protocols in Molecular Cloning*, Vols 1–3, John Wiley & Sons Inc., New York, NY .

Bard J.B.L., McConnell J. E. & Davies J. A. (1994) Towards a genetic basis for kidney development. *Mechanisms of Development* **48,** 3–11.

Bard J.B.L., Davies J.A., Karavanova I., Lehtonen E., Sariola H. & Vainio S. (1996) Kidney development: the inductive interactions. *Seminars in Cell and Developmental Biology* **7,** 195–202.

Barnes J.D., Crosby J.L., Jones C.M., Wright C.V.E. & Hogan B.L.M. (1994) Embryonic expression of *Lim-1*, the mouse homolog of Xenopus *Xlim-1*, suggests a role in lateral mesoderm differentiation and neurogenesis. *Developmental Biology* **161,** 168–178.

Beier D.R., Morton C.C., Leder A., Wallace R. & Leder P. (1989) Perinatal lethality (ple): a mutation caused by integration of a transgene into distal mouse chromosome 15. *Genomics* **4,** 498–504.

Benson R.C., Meyer R.A., Zaruba M.E. & McKann G.M. (1979) Cellular autofluorescence-is it due to flavins? *Journal of Histochemistry and Cytochemistry* **27,** 44–48.

Beullens M., Van Eynde A., Stalmans W. & Bollen M. (1992) The isolation of novel inhibitory polypeptides of protein phosphatase 1 from bovine thymus nuclei. *Journal of Biological Chemistry* **267,** 16 538–16 544.

Beullens M., Stalmans W. & Bollen M. (1996) Characterization of a ribosomal inhibitory polypeptide of protein phosphatase-1 from rat liver. *European Journal of Biochemistry* **239,** 183–189.

Bibb J.A., Yan Z., Svenningsson P., Snyder G.L., Pieribone V.A., Horiuchi A., Nairn A.C., Messer A. & Greengard P. (2000) Severe deficiencies in dopamine signaling in presymptomatic Huntington's disease mice. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 6809–6814.

Blake J.A., Eppig J.T., Richardson J.E., Bult C.J., Kadin J.A. & the Mouse Genome Database Group (2001) The Mouse Genome Database (MGD): Integration Nexus for the Laboratory Mouse. *Nucleic Acids Research* **29,** 91-94.

Breier G., Albrecht U., Sterrer S. & Risau W. (1992) Expression of vascular endothelial growth-factor during embryonic angiogenesis and endothelial-cell differentiation. *Development* **114,** 521–532.

Buehr, M., Gu S. & McLaren A. (1993) Mesonephric contribution to testis differentiation in the fetal mouse. *Development* **117,** 273–281.

Bullock S.L., Fletcher J.M., Beddington R.S.P. & Wilson V.A. (1998) Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes and Development* **12,** 1894–1906.

Capel B., Albrecht K.H., Washburn L.L. & Eicher E.M. (1999) Migration of mesonephric cells into the mammalian gonad depends on *Sry*. *Mechanisms of Development* **84,** 127–131.

Carmeliet P., Ferreira V., Breier G., Pollefeyt S., Kieckens L., Gertsenstein M., Fahrig M., Vandenhoeck A., Harpal K., Eberhardt C., Declercq C., Pawling J., Moons L., Collen D., Risau W. & Nagy A. (1996) Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature* **380,** 435–439.

Celli G., LaRochelle W.J., Mackem S., Sharp R. & Merlino G. (1998) Soluble dominant-negative receptor uncovers essential roles for fibroblast growth factors in multi-organ induction and patterning. *EMBO Journal* **17,** 1642–1655.

Chalfie M., Tu Y., Euskirchen G., Ward W.W. & Prasher D.C. (1994) Green fluorescent protein as a marker for gene expression. *Science* **263,** 802–805.

Chowdhury K., Bonaldo P., Torres M., Stoykova A. & Gruss P. (1997) Evidence for the stochastic integration of gene trap vectors into the mouse germline. *Nucleic Acids Research* **25,** 1531–1536.

Clover J., Oates J. & Edwards C. (1997) Anti-cytokeratin 5/6: a positive marker for epithelioid mesothelioma. *Histopathology* **31,** 140–143.

Compton J.G., Ferrara D.M., Yu D.W., Recca V., Freedberg I.M. & Bertolino, A.P. (1991) Chromosomal localization of mouse hair keratin genes. *Annals of the New York Academy of Sciences* **642,** 32–43.

Connor J.H., Quan H.N., Ramaswamy N.T., Zhang L.F., Barik S., Zheng J.H., Cannon J.F., Lee E.Y.C. & Shenolikar S. (1998) Inhibitor-1 interaction domain that mediates the inhibition of protein phosphatase-1. *Journal of Biological Chemistry* **273,** 27 716–27 724.

Connor J.H., Frederick D., Huang H., Yang J., Helps N.R., Cohen P.T.W, Nairn A.C., DePaoli-Roach A., Tatchell K. & Shenolikar S. (2000) Cellular mechanisms regulating protein phosphatase-1: a key functional interaction between inhibitor-2 and the type 1 protein phosphatase catalytic subunit. *Journal of Biological Chemistry* **275,** 18 670–18 675.

Davies J.A. & Garrod D.R. (1995) Induction of early stages of kidney tubule differentiation by lithium ions. *Developmental Biology* **167,** 50–60.

Davies J., Lyon M., Gallagher J. & Garrod D. (1995) Sulfated proteoglycan is required for collecting duct growth and branching but not nephron formation during kidney development. *Development* **121,** 1507–1517.

Davies J.A. (1996) Mesenchyme to epithelium transition during development of the mammalian kidney tubule. *Acta Anatomica* **156,** 187–210.

Davies J.A. & Brandli A.W. (1997) The Kidney Development Database http://golgi.ana.ed.ac.uk/kidhome.html.

Davies J.A. & Bard J.B.L. (1998) The development of the kidney. *Current Topics in Developmental Biology* **39,** 245–301.

Doyle K. (1996) Protocols and Applications Guide: The Source of Discovery. 3rd edn. Promega Corporation, Madison, WI.

Dressler G.R., Deutsch U., Chowdhury K., Nornes H.O. & Gruss P. (1990) *Pax2*, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* **109,** 787–795.

Dressler G.R. & Douglass E.C. (1992) Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms Tumour. *Proceedings of the National Academy of Sciences of the United States of America* **89,** 1179–1183.

Durbec P., Marcos-Gutierrez C.V., Kilkenny C., Grigoriou M., Wartiowaara K., Suvanto P., Smith D., Ponder B., Costantini F., Saarma M., Sariola H. & Pachnis V. (1996) GDNF signalling through the Ret receptor tyrosine kinase. *Nature* **381,** 789–793.

Ekblom P. (1981) Formation of basement-membranes in the embryonic kidney – an immunohistological study. *Journal of Cell Biology* **91,** 1–10.

Ekblom P., Klein G., Ekblom M. & Sorokin L. (1991) Laminin isoforms and their receptors in the developing kidney. *American Journal of Kidney Diseases* **17,** 603–605.

Ekblom P. (1996) Receptors for laminins during epithelial morphogenesis. *Current Opinion in Cell Biology* **8,** 700–706.

Elbrecht A., DiRenzo J., Smith R.G. & Shenolikar S. (1990) Molecular cloning of protein phosphatase inhibitor-1 and its expression in rat and rabbit tissues. *Journal of Biological Chemistry* **265,** 13 415–13 418.

Endo S., Zhou X., Connor J., Wang B. & Shenolikar S. (1996) Multiple strucural elements define the specificity of recombinant human inhibitor-1 as a protein phosphatase-1 inhibitor. *Biochemistry* **35,** 5220–5228.

Eto M., Ohmuri T., Suzuki M., Furuya K. & Morita F. (1995) A novel protein phosphatase-1 inhibitory protein potentiated by protein kinase C. Isolation from porcine aorta media and characterization. *Journal of Biochemistry* **118,** 1104–1107.

Eto, M., Senba S., Morita F. & Yazawa M. (1997) Molecular cloning of a novel phosphorylation-dependent inhibitory protein of protein phosphatase-1 (CPI17) in smooth muscle: its specific localization in smooth muscle. *FEBS Letters* **410,** 356–360.

Fantes J.A., Bickmore W.A., Fletcher J.M., Ballesta F., Hanson I.M. & Vanheyningen V. (1992) Submicroscopic deletions at the wagr locus, revealed by nonradioactive insitu hybridization. *American Journal of Human Genetics* **51,** 1286–1294.

Ferrara N., Carver-Moore K., Chen H., Dowd M., Lu L., O'Shea K.S., Powell-Braxton L., Hillan K.J. & Moore M.W. (1996) Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature* **380,** 439–442.

Fine L.G. (1996) Gene transfer into the kidney: promise for unravelling disease mechanisms, limitations for human gene therapy. *Kidney International* **49,** 612–619.

Fouser L. & Avner E.D. (1993) Normal and abnormal nephrogenesis. *American Journal of Kidney Diseases* **21,** 64–70.

Fryckstedt J., Aperia A., Snyder G. & Meister B. (1993) Distribution of dopamine- and cAMP-dependent phosphoprotein (DARPP-32) in the developing and mature kidney. *Kidney International* **44,** 495–502.

Fujii T., Pichel J.G., Taira M., Toyama R., Dawid I.B. & Westphal H. (1994) Expression patterns of the murine lim class homeobox gene lim1 in the developing brain and excretory system. *Developmental Dynamics* **199,** 73–83.

Gardiner-Garden M. & Frommer M. (1987) CpG islands in vertebrate genomes. *Journal of Molecular Biology* **196,** 261–282.

Gieseler R., Hoffmann, P.R., Kuhn R., Fayyazi A., Stojanovic T., Schlemminger R. & Peters J.H. (1997) Enrichment and characterization of dendritic cells from rat renal mesangium. *Scandinavian Journal of Immunology* **46,** 587–596.

Grobstein C. (1956) Trans-filter induction of tubules in mouse metanephric mesenchyme. *Experimental Cell Research* **10,** 424–444.

Gupta R.C., Neumann J., Watanabe A.M., Lesch M. & Sabbah H.N. (1996) Evidence for presence and hormonal regulation of protein phosphatase inhibitor-1 in ventricular cardiomyocyte. *American Journal of Physiology* **270,** H1159–H1164.

Hacker A., Capel B., Goodfellow P. & Lovell-Badge R. (1995) Expression of *Sry*, the mouse sex determining gene. *Development* **121,** 1603–1614.

Hatini V., Huh S.O., Herzlinger D., Soares V.C. & Lai E. (1996) Essential role of stromal mesenchyme in kidney morphogenesis revealed by targeted disruption of Winged Helix transcription factor *BF-2*. *Genes and Development* **10,** 1467–1478.

Hay E.D. (1995) An overview of epithelio-mesenchymal transformation. *Acta Anatomica* **154,** 8–20.

Heine U.I., Flanders K., Roberts A.B., Munoz E.F. & Sporn M.B. (1987) Immunocytochemical localization of TGF-beta during embryonal development in mice. *Proceedings of the American Association for Cancer Research* **28,** 53–53.

Hemmings H. C., Girault J., Nairn A., Bertuzzi G. & Greengard P. (1992) Distribution of protein phosphatase inhibitor-1 in brain and peripheral tissues of various species: comparison with DARPP-32. *Journal of Neurochemistry* **59,** 1053–1061.

Herzlinger D. (1994) Renal stem cells and the lineage of the nephron. *Annual Review of Physiology* **56,** 671–689.

Huang H.B., Horiuchi A., Watanabe T., Shih S.R., Tsay H.J., Li H.C., Greengard P. & Nairn A.C. (1999) Characterization of the inhibition of protein phosphatase-1 by DARPP-32 and inhibitor-2. *Journal of Biological Chemistry* **274,** 7870–7878.

Huang K.X. & Paudel H.K. (2000) Ser(67)-phosphorylated inhibitor 1 is a potent protein phosphatase 1 inhibitor. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 5824–5829.

Kakinoki Y., Somers J. & Brautigan D.L. (1997) Multisite phosphorylation and the nuclear localization of phosphatase Inhibitor 2-green fluorescent protein fusion

protein during S phase of the cell growth cycle. *Journal of Biological Chemistry* **272,** 32 308–32 314.

Kallapur S., Ormsby I. & Doetschman T. (1999) Strain dependency of TGF beta 1 function during embryogenesis. *Molecular Reproduction and Development* **52,** 341–349.

Karavanov A.A., Karavanova I., Perantoni A. & Dawid I.B. (1998) Expression pattern of the rat Lim-1 homeobox gene suggests a dual role during kidney development. *International Journal of Developmental Biology* **42,** 61–66.

Karavanova I.D., Dove L.F., Resau J.H. & Perantoni A.O. (1996) Conditioned medium from a rat ureteric bud cell line in combination with bFGF induces complete differentiation of isolated metanephric mesenchyme. *Development* **122,** 4159–4167.

Kaufman M.H. & Bard J.B.L. (1999) *The Anatomical Basis of Mouse Development*. Academic Press, San Diego, CA.

Kispert A., Vainio S., Shen L., Rowitch D.H. & McMahon A.P. (1996) Proteoglycans are required for maintenance of *Wnt-11* expression in the ureter tips. *Development* **122,** 3627–3637.

Kispert A., Vainio S. & McMahon A.P. (1998) Wnt-4 is a mesenchymal signal for epithelial transformation of metanephric mesenchyme in the developing kidney. *Development* **125,** 4225–4234.

Kitamoto Y., Tokunaga H. & Tomita K. (1997) Vascular endothelial growth factor is an essential molecule for mouse kidney development: glomerulogenesis and nephrogenesis. *Journal of Clinical Investigation* **99,** 2351–2357.

Kjelsberg C., Sakurai H., Spokes K., Birchmeier C., Drummond I., Nigam S. & Cantley L.G. (1997) Met –/– kidneys express epithelial cells that chemotax and form tubules in response to EGF receptor ligands. *American Journal of Physiology* **272,** F222–F228.

Klein P.S. & Melton D.A. (1996) A molecular mechanism for the effect of lithium on development. *Proceedings of the National Academy of Sciences of the United States of America* **93,** 8455–8459.

Korhonen M., Ylanne J., Laitinen L. & Virtanen I. (1990) The alpha-1-alpha-6 subunits of integrins are characteristically expressed in distinct segments of developing and adult human nephron. *Journal of Cell Biology* **111,** 1245–1254.

Koseki C., Herzlinger D & Al-Awqati (1992) Apoptosis in metanephric development. *Journal of Cell Biology* **119,** 1327–1333.

Kreidberg J.A., Sariola H., Loring J.M., Maeda M., Pelletier J., Housman D. & Jaenisch R. (1993) WT-1 is required for early kidney development. *Cell* **74,** 679–691.

Kreidberg J.A., Donovan M.J., Goldstein S.L., Rennke H., Shepherd K., Jones R.C. & Jaenisch R. (1996) Alpha 3 beta 1 integrin has a crucial role in kidney and lung organogenesis. *Development* **122,** 3537–3547.

Kume T., Deng K.Y. & Hogan B.L.M. (2000) Murine forkhead/winged helix genes Foxc1 (Mf1) and Foxc2 (Mfh1) are required for the early organogenesis of the kidney and urinary tract. *Development* **127,** 1387–1395.

Kuure S., Vuolteenaho R. & Vainio S. (2000) Kidney morphogenesis: cellular and molecular regulation. *Mechanisms of Development* **92,** 31–45.

Lennon G., Auffray C., Polymeropoulos M. & Soares M.B. (1996) The IMAGE consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33,** 151–152.

Lehnert S.A. & Akhurst R.J. (1988) Embryonic expression pattern of TGF-beta type-1 RNA suggests both paracrine and autocrine mechanisms of action. *Development* **104,** 263–273.

Leonard M., Ryan M.P., Watson A.J., Schramek H. & Healy E. (1999) Role of MAP kinase pathways in mediating IL-6 production in human primary mesangial and proximal tubular cells. *Kidney International* **56,** 1366–1377.

Leveen P., Pekny M., Gebremedhin S., Swolin B., Larsson E. & Betsholtz C. (1994) Mice deficient for PDGF-B show renal, cardiovascular, and hematological abnormalities. *Genes and Development* **8,** 1875–1887.

Lin L.F.H., Doherty D.H., Lile J.D., Bektesh S. & Collins F. (1993) GDNF – a glial-cell line derived neurotrophic factor for midbrain dopaminergic-neurons. *Science* **260,** 1130–1132.

Lindahl P., Hellstrom M., Kalen M., Karlsson L., Pekny M., Pekna M., Soriano P. & Betsholtz C. (1998) Paracrine PDGF-B/PDGF-R beta signaling controls mesangial cell development in kidney glomeruli. *Development* **125,** 3313–3322.

Liu L. & Barajas L. (1993) The rat renal nerves during development. *Anatomy and Embryology* **188,** 345–361.

Lowenstein P.R., Shering A.F., MacDougall L.K. & Cohen P. (1995) Immunolocalisation of protein phosphatase inhibitor-1 in the cerebral cortex of the rat, cat and ferret. *Brain Research* **676,** 80–92.

MacDougall L. K., Campbell D.G., Hubbar M.J. & Cohen P. (1989) Partial structure and hormaonal regulation of rabbit liver inhibitor-1; distribution of inhibitor-1 and inhibitor-2 in rabbit and rat tissues. *Biochemica et Biophysica Acta* **1010,** 218–226.

Malek A. & Khaledi M. (1999) Expression and analysis of green fluorescent proteins in human embryonic kidney cells by capillary electrophoresis. *Analytical Biochemistry* **268,** 262–269.

Martineau J., Nordqvist K., Tilmann C., Lovell-Badge R. & Capel B. (1997) Male-specific cell migration into the developing gonad. *Current Biology* **7**, 958–968.

McLaren L., Boyle S., Mason J.O. & Bard J.B.L. (2000) Expression and genomic characterization of protein phosphatase inhibitor-1: a novel marker for mesothelium in the mouse. *Mechanisms of Development* **96,** 237–241.

Meister B., Fryckstedt J., Schalling M., Cortes R., Hokfelt T., Aperia A., Hemmings H.C., Nairn A.C., Ehrlich M. & Greengard P. (1989) Dopamine-regulated and camp-regulated phosphoprotein (DARPP-32) and dopamine DA_1 agonist-sensitive Na^+, K^+-ATPase in renal tubule cells. *Proceedings of the National Academy of Sciences of The United States of America* **86,** 8068–8072.

Miner J.H. & Li C. (2000) Defective glomerulogenesis in the absence of laminin alpha 5 demonstrates a developmental role for the kidney glomerular basement membrane. *Developmental Biology* **217,** 278–289.

Moore M.W., Klein R.D., Farinas I., Sauer H., Armanini M., Phillips H., Reichardt L.F., Ryan A.M., Carver-Moore K. & Rosenthal A. (1996) Renal and neuronal abnormalities in mice lacking GDNF. *Nature* **382,** 76–79.

Mountford P.S. & Smith A.G. (1995) Internal ribosome entry sites and dicistronic RNAs in mammalian transgenesis. *Trends in Genetics* **11,** 179–184.

Muller U., Wang D., Denda S., Meneses J.J., Pedersen R.A. & Reichardt L.F. (1997) Integrin $\alpha 8\beta 1$ is critically important for epithelial-mesenchymal interactions during kidney morphogenesis. *Cell* **88,** 603–613.

Mumby M.C. & Walter G. (1993) Protein serine/threonine phosphatases: structure, regulation and functions in cell growth. *Physiological Reviews* **73,** 673–699.

Oates J. & Edwards C. (2000) HBME-1, MOC-31, WT1 and calretinin: an assessment of recently described markers for mesothelioma and adenocarcinoma. *Histopathology* **36,** 341–347.

Oliver C.J. & Shenolikar S. (1998) Physiologic importance of protein phosphatase inhibitors. *Frontiers in Bioscience* **3,** 961–972.

Orth S.R., Ritz E. & Suter-Crazzolara C. (2000) Glial cell line-derived neurotrophic factor (GDNF) is expressed in the human kidney and is a growth factor for human mesangial cells. *Nephrology Dialysis Transplantation* **15,** 589–595.

Osawa Y., Nakagama H., Shima H., Sugimura T. & Nagao M. (1996) Identification and characterization of three isotypes of protein phosphatase inhibitor-2 and their expression profiles during testis maturation in rats. *European Journal of Biochemistry* **242,** 793–798.

Pepicelli C.V., Kispert A., Rowitch D.H. & McMahon A.P. (1997) GDNF induces branching and increased cell proliferation in the ureter of the mouse. *Developmental Biology* **192,** 193–198.

Pichel J.G., Shen, L.Y., Sheng H.Z., Granholm A.C., Drago J., Grinberg A., Lee E.J., Huang S.P., Saarma M., Hoffer B.J., Sariola H. & Westphal H. (1996) Defects in enteric innervation and kidney development in mice lacking GDNF. *Nature* **382,** 73–76.

Piscione T.D. & Rosenblum N.D. (1999) The malformed kidney: disruption of glomerular and tubular development. *Clinical Genetics* **56,** 341–356.

Plisov S.Y., Yoshino K., Dove L.F., Higinbotham K.G. & Rubin J.S. (2001) TGFβ2, LIF and FGF2 cooperate to induce nephrogenesis. *Development* **128,** 1045–1057.

Pritchard-Jones K., Fleming S., Davidson D., Bickmore W., Porteous D., Gosden C., Bard J., Buckler A., Pelletier J., Housman D., van Heyningen V. & Hastie N. (1990) The candidate Wilm's tumor gene is involved in genitourinary development. *Nature* **346,** 194–197.

Qiao J., Sakurai H. & Nigam S.K. (1999a) Branching morphogenesis independent of mesenchymal-epithelial contact in the developing kidney. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 7330–7335.

Qiao J.Z., Uzzo R., Obara-Ishihara T., Degenstein L., Fuchs E. & Herzlinger D. (1999b) FGF-7 modulates ureteric bud growth and nephron number in the developing kidney. *Development* **126,** 547–554.

Rahilly M.A. & Fleming S. (1992) Differential expression of integrin alpha-chains by renal epithelial-cells. *Journal of Pathology* **167,** 327–334

Richards W.G., Sweeney W.E., Yoder B.K., Wilkinson J.E., Woychik R.P. & Avner E.D. (1998) Epidermal growth factor receptor activity mediates renal cyst formation in polycystic kidney disease. *Journal Clinical Investigation* **101,** 935–939.

Risau W. & Ekblom P. (1986) Production of a heparin-binding angiogenesis factor by the embryonic kidney. *Journal of Cell Biology* **103,** 1101–1107.

Robillard J.E., Guillery E.N., Segar J.L., Merrill D.C. & Jose P.A. (1993) Influence of renal nerves on renal-function during development. *Pediatric Nephrology* **7,** 667–671.

Rogers S.A., Ryan G., Purchio A.F. & Hammerman M.R. (1993) Metanephric transforming growth-factor-beta-1 regulates nephrogenesis *in vitro*. *American Journal of Physiology* **264,** F996–F1002.

Rosenblum N.D. & Yager T.D. (1997) Changing patterns of gene expression in developing mouse kidney, as probed by differential mRNA display combined with cDNA library screening. *Kidney International* **51,** 920–925.

Sainio K., Suvanto P., Davies J., Wartiovaara, J., Wartiovaara K., Saarma M., Arumae U., Meng X., Lindahl M., Pachnis V. & Sariola H. (1997) Glial-cell-line-derived neurotrophic factor is required for bud initiation from ureteric epithelium. *Development* **124,** 4077–4087.

Sakagami H., Ebina K. & Kondo H. (1994) Localization of phosphatase inhibitor-1 mRNA in the developing and adult rat brain in comparison with that of protein phosphatase-1 mRNAs. *Molecular Brain Research* **25,** 7–18.

Sakurai H. & Nigam S.K. (1997) Transforming growth factor-beta selectively inhibits branching morphogenesis but not tubulogenesis. *American Journal of Physiology* **272,** F139–146.

Sakurai H., Tsukamoto T., Kjelsberg C.A., Cantley L.G. & Nigam S.K. (1997) EGF receptor ligands are a large fraction of in vitro branching morphogens secreted by embryonic kidney. *American Journal of Physiology* **273,** F463–F472.

Sambrook J., Fritsch E.F. & Maniatas T. (eds, 1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NJ.

Sanchez M.P., Silos-Santiago I., Frisen J., He B., Lira S.A. & Barbacid M. (1996) Renal agenesis and the absence of enteric neurons in mice lacking GDNF. *Nature* **382,** 70–73.

Sanyanusin P., Schimmenti L.A., McNoe L.A., Ward T.A., Pierpont M.E.M., Sullivan M.J., Dobyns W.B. & Eccles M.R. (1995) Mutation of the pax2 gene in a family with optic-nerve colobomas, renal anomalies and vesicoureteral reflux. *Nature Genetics* **9,** 358–364.

Sariola H., Timpl R., Vondermark K., Mayne R., Fitch J.M., Linsenmayer T.F., Ekblom P. (1984) Dual origin of glomerular basement-membrane. *Developmental Biology* **101,** 86–96.

Sastry S.K. & Horwitz A.F. (1996) Adhesion-growth factor interactions during differentiation: An integrated biological response. *Developmental Biology* **180,** 455–467.

Saxen L., Salonen J., Ekblom P. & Nordling S. (1983) DNA-synthesis and cell generation cycle during determination and differentiation of the metanephric mesenchyme. *Developmental Biology* **98,** 130–138.

Schuchardt A., Dagati V., Larssonblomberg L., Costantini F. & Pachnis V. (1994) Defects in the kidney and enteric nervous-system of mice lacking the tyrosine kinase receptor Ret. *Nature* **367,** 380–383.

Schuchardt A., D'Agati V., Pachnis V. & Costantini F. (1996) Renal agenesis and hypodysplasia in *ret-k⁻* mutant mice result from defects in ureteric bud development. *Development* **122,** 1919–1929.

Scrimgeour A.G., Allen P.B., Fienberg A.A., Greengard P. & Lawrence J.C., Jr (1999) Inhibitor-1 is not required for the activation of glycogen synthase by insulin in skeletal muscle. *Journal of Biological Chemistry* **274,** 20 949–20 952.

Senba S., Eto M. & Yazawa M. (1999) Identification of trimeric myosin phosphatase (PP1M) as a target for a novel PKC-potentiated protein phosphatase-1 inhibitory protein (CPI17) in porcine aorta smooth muscle. *Journal of Biochemistry* **125,** 354–362.

Shawlot W. & Behringer R.R. (1995) Requirement for Lim1 in head-organizer function. *Nature* **374,** 425–430.

Shirato H., Shima H., Sakashita G., Nakano T., Ito M., Lee E.Y.C. & Kikuchi K. (2000) Identification and characterization of a novel protein inhibitor of type 1 protein phosphatase. *Biochemistry* **39,** 13 848–13 855.

Shull M.M., Ormsby I., Kier A.B., Pawlowski S., Diebold R.J., Yin M.Y., Allen R., Sidman C., Proetzel G., Calvin D., Annunziati N. & Doetschman T. (1992) Targeted disruption of the mouse transforming growth factor-beta-1 gene results in multifocal inflammatory disease. *Nature* **359,** 693–699.

Sonnenberg E., Meyer D., Weidner K.M., Birchmeier C. (1993) Scatter factor/hepatocyte growth-factor and its receptor, the c-met tyrosine kinase, can mediate a signal exchange between mesenchyme and epithelia during mouse development. *Journal of Cell Biology* **123,** 223-235

Soriano P. (1994) Abnormal kidney development and hematological disorders in PDGF beta-Receptor mutant mice. *Genes and Development* **8,** 1888–1896.

Sorokin L., Sonnenberg A., Aumailley M., Timpl R. & Ekblom P. (1990) Recognition of the laminin E8 cell-binding site by an integrin possessing the α6 subunit is essential for epithelial polarization in developing kidney tubule. *Journal of Cell Biology* **111,** 1265–1273.

Sorokin L. & Ekblom P. (1992) Development of tubular and glomerular cells of the kidney. *Kidney International* **41,** 657–664.

Stark K., Vainio S., Vassileva G. & McMahon A.P. (1994) Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by *Wnt-4*. *Nature* **372,** 679–683.

Stralfors P., Hemmings H.C. & Greengard P. (1989) Inhibitors of protein phosphatase-1: inhibitor-1 of bovine adipose tissue and dopamine- and cAMP-regulated phosphoprotein of bovine brain are identical. *European Journal of Biochemistry* **180,** 143–148.

Stricklett P.K., Nelson R.D. & Kohan D.E. (1999) The Cre loxP system and gene targeting in the kidney. *American Journal of Physiology* **276,** F651–F657.

Svennilson J., Durbeej M., Celsi G., Laestadius A., da Cruz e Silva E.F., Ekblom P. & Aperia A. (1995) Evidence for a role of protein phosphatases 1 and 2A during early nephrogenesis. *Kidney International* **48,** 103–110.

Torres M., Gomez-Pardo E., Dressler G.R. & Gruss P. (1995) *Pax-2* controls multiple steps of urogenital development. *Development* **121,** 4057–4065.

Van Eynde A., Wera S., Beullens M., Torrekens S., Van Leuven F., Stalmans W. & Bollen M. (1995) Molecular cloning of NIPP-1, a nuclear inhibitor of protein phosphatase-1, reveals homology with polypeptides involved in RNA processing. *Journal of Biological Chemistry* **270,** 28 068–28 074.

Vize P.D., Seufert D.W., Carroll T.J. & Wallingford J.B. (1997) Model systems for the study of kidney development: use of the pronephros in the analysis of organ induction and patterning. *Developmental Biology* **188,** 189–204.

Walaas S.I., Aswad D.W. & Greengard P. (1983) A dopamine- and AMP-regulated phosphoprotein enriched indopamine-innervated brain regions. *Nature* **301,** 69–71.

Wartenberg H., Kinsky I., Viebahn C. & Schmolke C. (1991) Fine structural characteristics of testicular cord formation in the developing rabbit gonad. *Journal of Electron Microscopy Technique* **19,** 133–157.

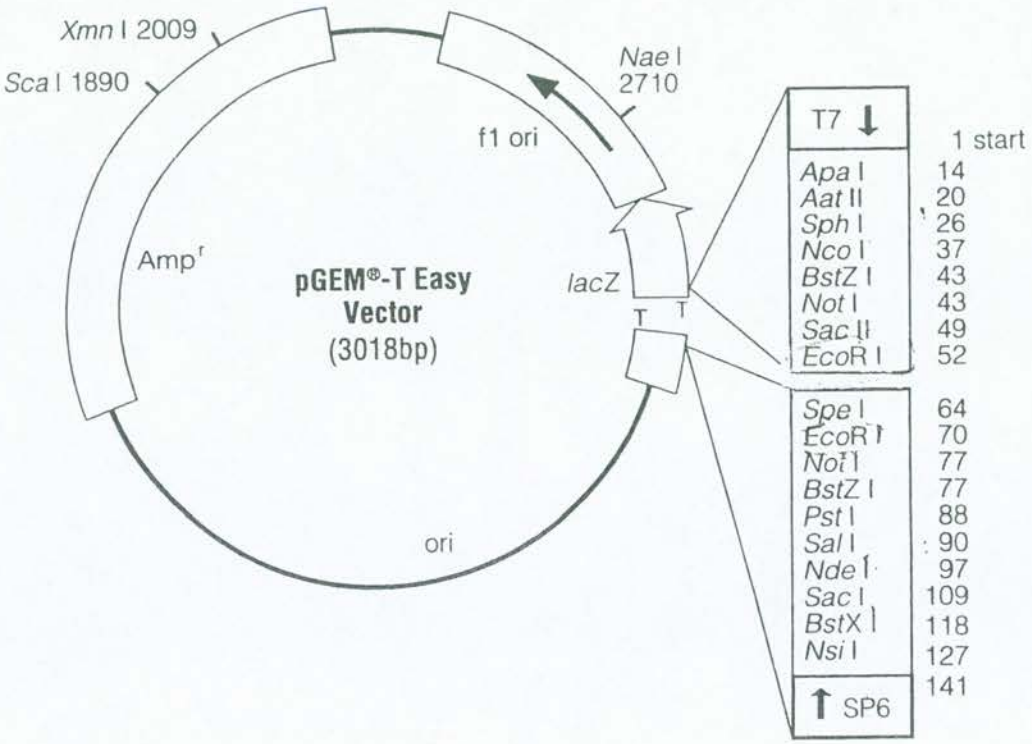
Wera, S., Van Eynde A., Stalmans W. & Bollen M. (1997) Inhibition of translation by mRNA encoding NIPP-1, a nuclear inhibitor of protein phosphatase-1. *European Journal of Biochemistry* **247,** 411–415.

Whitaker D. (2000) The cytology of malignant mesothelioma. *Cytopathology* **11,** 139–151.

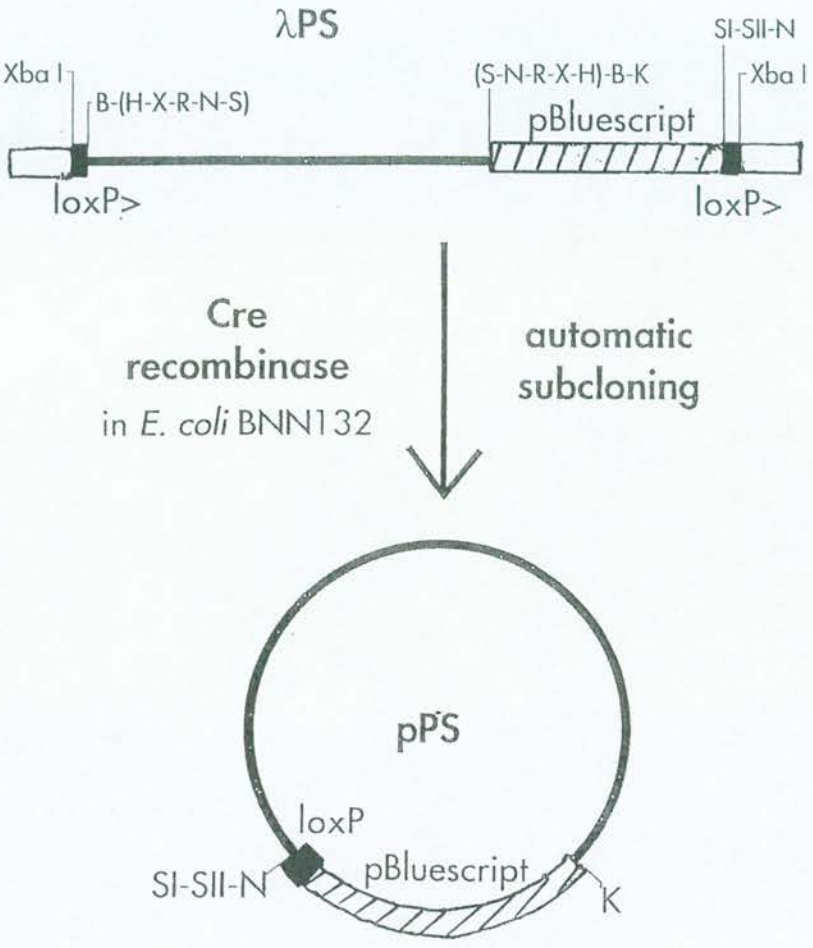
Williams P.L., Bannister L.H., Berry M.M., Collins P., Dyson M., Dussek J.E. & Ferguson M.W.J. (eds) (1995) Gray's Anatomy, 38th edn. Churchill Livingstone, Edinburgh.

Appendix 1

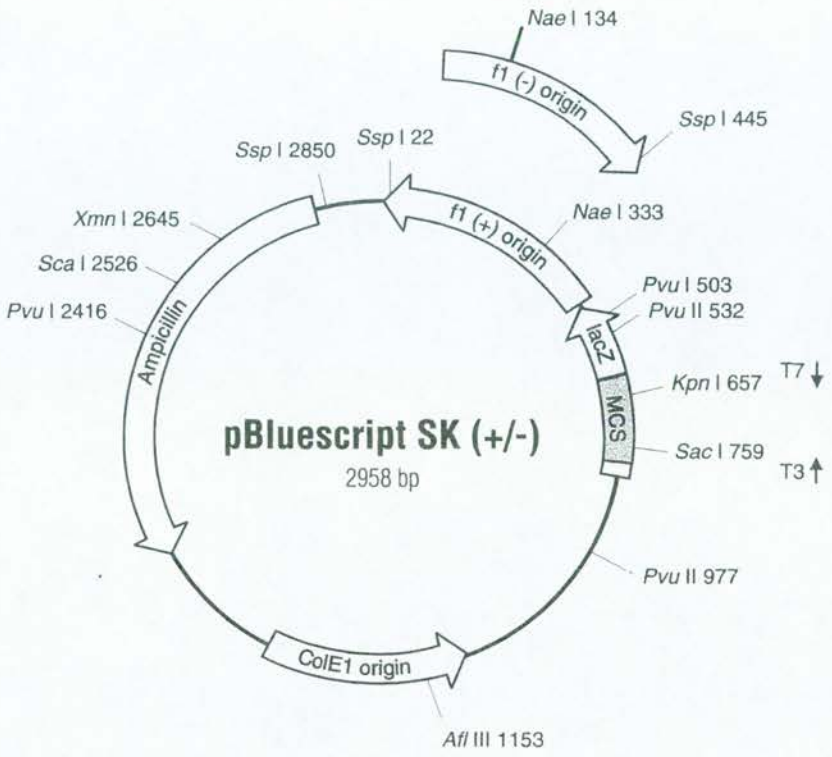
Figure A1. The maps of vectors used in the research project: (a) pGEM®-T Easy (Promega); (b) the λPS vector of the mouse genomic library (MoBiTec), showing the position of the loxP and restriction sites; (c) pBluescriptSK+/– (Stratagene); (d) pBluescriptIIKS+/– (Stratagene); (e) pGT1·8IRESβgeo (built by W. Skarnes, P. Mountford and P. Tate in 1992); (f) pEGFP-1 (Clontech); and (g) pEGFP-N1 (Clontech).

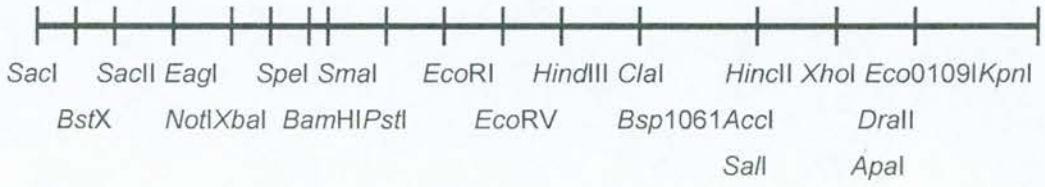


(a) The vector pGEM®-T Easy from Promega was used to clone the 336-bp PCR product from 1912-d05. The restriction sites in the multiple cloning site (MCS) are shown.

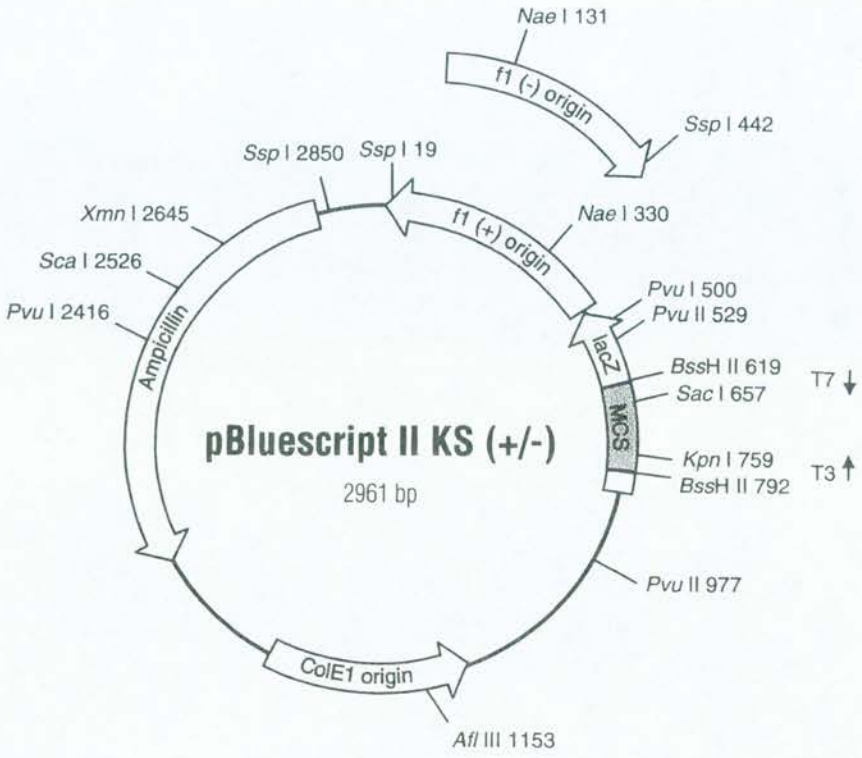


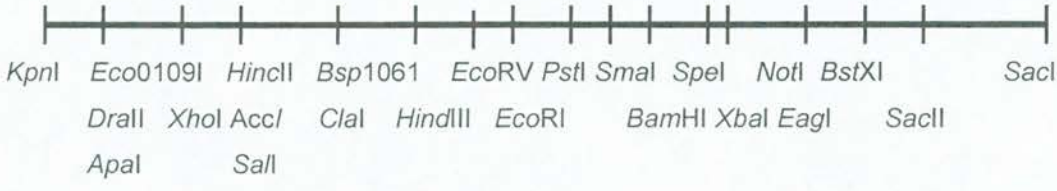
(b) The structure of the λPS vector and the resulting insert DNA, pBluescript and loxP sites after automatic subcloning in Cre recombinase-expressing cells. In the final plasmid, the sites *Hind*III (H), *Xho*I (X), *Eco*RI (R), *Not*I (N) and *Sal*I (S) are lost, leaving only *Sac*I (SI), *Sac*II (SII), *Not*I and *Kpn*I (K) in the pBluescript backbone.



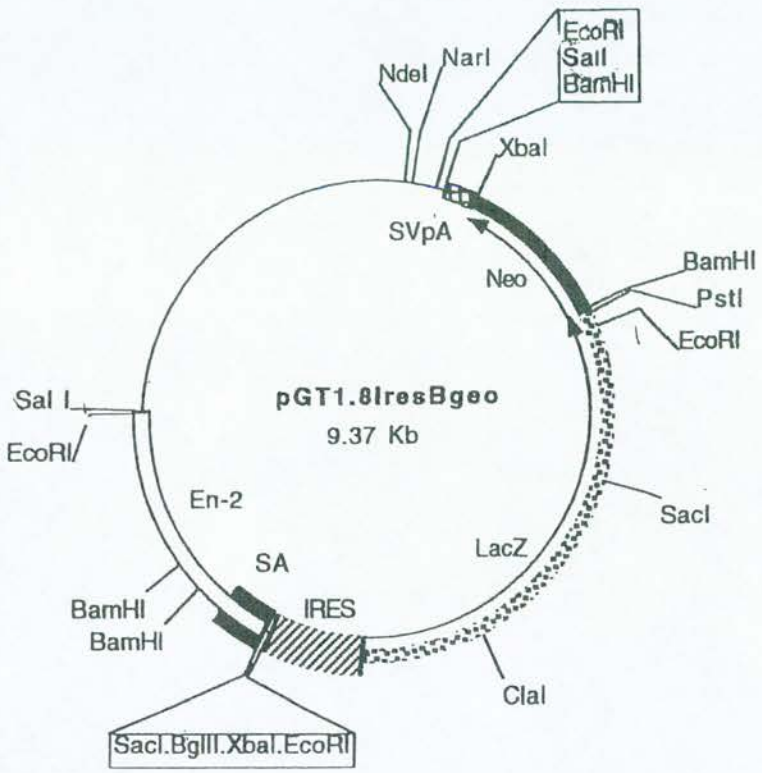


(c) pBluescriptSK (+/–) from Stratagene (above) is a 2958-bp phagemid derived from pUC19. The MCS shows the restriction sites present.

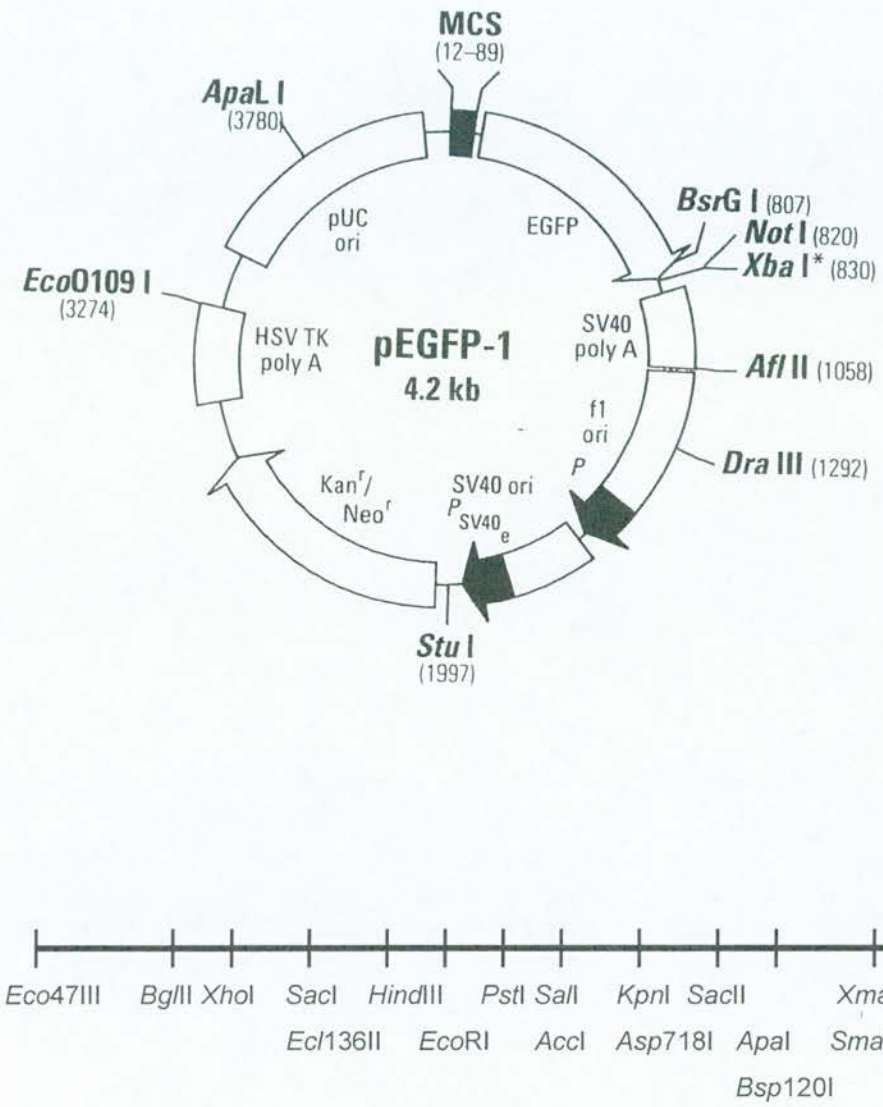




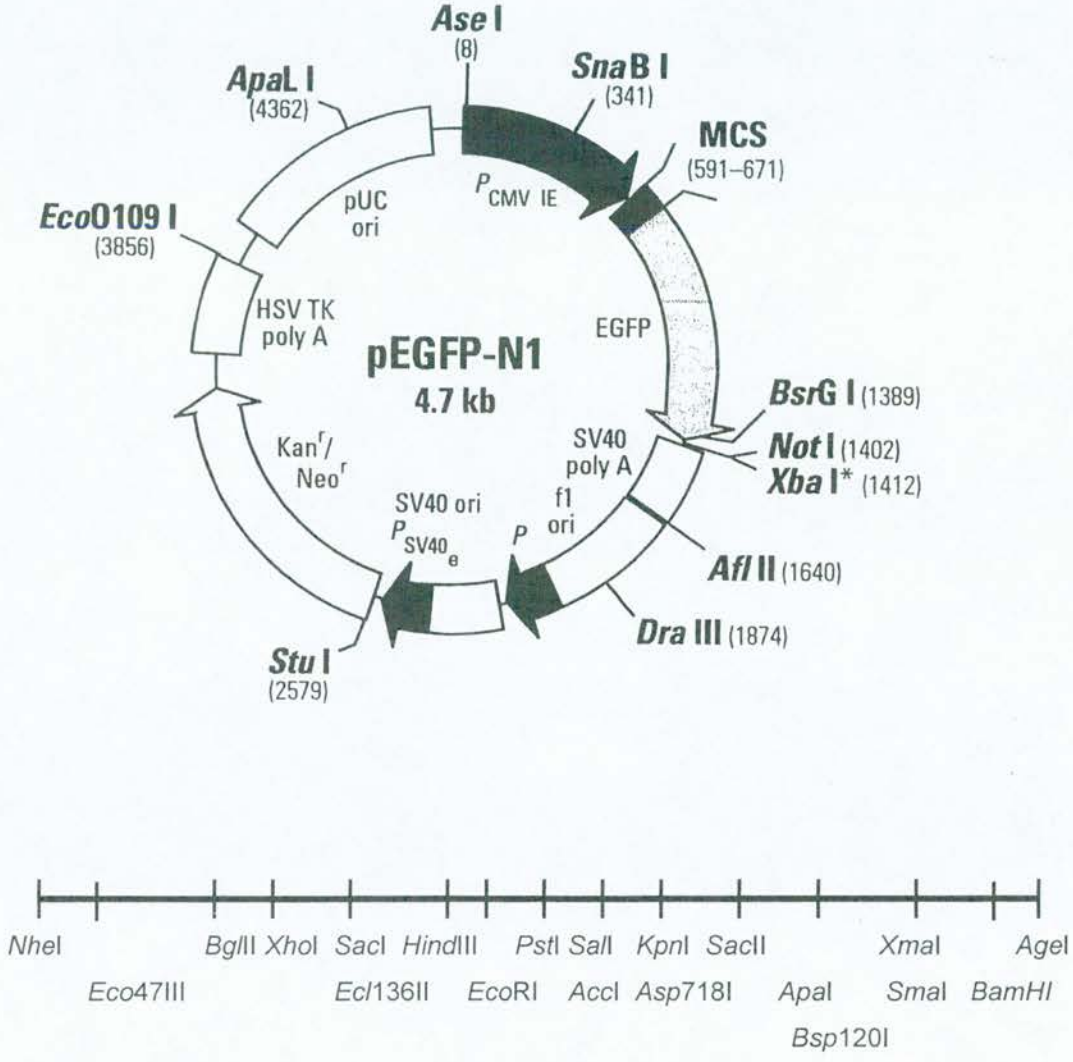
(d) pBluescriptIIKS from Stratagene (above) is a 2961-bp phagemid derived from pUC19. The MCS shows the restriction sites present.



(e) The vector pGT1·8IRESβgeo (above) is 9·37 kb and the main restriction sites are shown (built by W. Skarnes, P. Mountford and P. Tate in 1992).



(f) pEGFP-1 (above; Clontech; GenBank Accession number U55761) is 4·2 kb in size and encodes a red-shifted variant of wild-type GFP to give brighter fluorescence and better expression in mammalian cells. The MCS shows the restriction sites present.



(g) pEGFP-N1 (above; Clontech; GenBank Accession number U55762) is 4·7 kb in size and encodes for the same red-shifted variant of wild-type GFP as pEGFP-1. The MCS shows the restriction sites present.

Figure A2. Sequence analysis of 1912-d05. The mouse clone 1912-d05 was commercially sequenced (KC) and compared to the I.M.A.G.E. sequence (IM). Differences are indicated by highlighted pink nucleotides: (N) unknown nucleotide (yellow); and (–) a gap for best alignment.

```
     1                                                   50
KC   GTTCAGAGGCCAAGAATTCGGATCCATGGCCGCTGCCTGACCGGGAGCCA
IM                     ATTCGGATCCATGGCCGCTGCCTGACCGGGAGCCA
     51                                                  100
KC   TGGAGCCCGACAACAGCCCACGGAAGATCCAGTTTACGGTCCCGCTGCTG
IM   TGGAGCCCGACAACAGCCCACGGAAGATCCAGTTTACGGTCCCGCTGCTG
     101                                                 150
KC   GAGCCTCACCTGGACCCGGAGGCGGCGGA-GCAGATTCGGAGGCGCCGCCC
IM   GAGCCTCACCTGGACCCGGAG-CGGCGGACG-AGATTCGGAGGCGCCGCCC
     151                                                 200
KC   CACCCCTGCCACACTTGTGCTGACCAGTGATCAGTCCTCCCCAGAAATAG
IM   -ACCCCTGCCACACTTGTGCTGACCAGTGATCAGTCCTCCCCAGAAATAG
     201                                                 250
KC   ATGAAGACCGGATCCCCAACTCACTTCTCAAGTCCACCTTGTCAATGTCT
IM   ATGAAGACCGGATCCCCAACTCACTTCTCAAGTCCACCTTGTCAATGTCT
     251                                                 300
KC   CCACGGCAACGGAAGAAGATGACAAGGACCACACCCACCATGAAAGAGCT
IM   CCACGGCAACGGAAGAAGATGACAAGGACCACACCCACCATGAAAGAGCT
     301                                                 350
KC   CCAGACGATGGTTGAACATCACCTAGGGCAACAGAAGCAAGGGGAAAGAA
IM   CCAGACGATGGTTGAACATCACCTAGGGCAACAGAAGCAAGGGGAAGAA-
     351                                                 400
KC   CCTGAGGGAGCCACTGAGAGCACAGGGAACCAGGAGTCCTGCCCACCTGG
IM   CCTGAGGGAGCCACTGAGAGCACAGGGAACCAGGAGTCCTGCCCACCTGG
     401                                                 450
KC   GATTCCCAGACACAGGCTCNGCGTCCAAGGCCCAGATACCCCCNGGAACA
IM   -ATCCC-AGACACAGGCTC-GCGT-CAAGG-CCAGATACCCCCGGGA-CA
     451                                                     500
KC   GCACAAA-GTTCTGCAGAATCCAACCCCA-GACTCCAGGA-CA-TGTTGTGTTG
IM   GCACAAAAGT-CTGCAGAATCCAACCCCAAGACTC-AGGAGCAGTGTGGTGT-G
```

```
     501                                                        550
KC   GAGCCCAGA-CAGAG-ATTCTTCAGCCCCCATGCTACCATTG-ATTCCCAGGA
IM   GAGCCCAGAACAGAGGATTCTTCAGCCCACATGCTACCACTGGATTCCCAGGA
     551                                                        600
KC   CCA-GCTTGGTCTGACAATTTGTATCGGGGTTNNCCCNGCN TNTTGAATTC
IM   GCAAGC
```

Figure A3. Sequence analysis of 3141-o09. The clone was commercially sequenced (KC) and compared to the I.M.A.G.E. sequence (IM). The first 500 nucleotides are shown, but sequence homology differs after nucleotide 410. Differences are indicated by highlighted pink nucleotides: (N) unknown nucleotide (yellow); and (–) a gap for best alignment.

```
     1                                                   50
KC   GAGGCCAAGAATTCGGCACGAGGGAGGGAGCGCGCCCGCCCCAGCCCGAG
     51                                                  100
KC   GTCCGCCGCTGCCTGACCGGGAGCCATGGAGCCCGACAACAGCCCACGGA
IM    TCCGCCGCTGCCTGACCGGGAGCCATGGAGCCCGACAACAGCCCACGGA
     101                                                 150
KC   AGATCCAGTTTACGGTCCCGCTGCTGGAGCCTCACCTGGACCCGGAGGCG
IM   AGATCCAGTTTACGGTCCCGCTGCTGGAGCCTCACCTGGACCCGGAG-CG
     151                                                 200
KC   GCGGAGCAGATTCGGAGGCGCCGCCCCACCCCTGCCACACTTGTGCTGAC
IM   GCGGAA-NGATTCGGAGGCGCCGCCC-ACCCCTGCCACACTTGTGCTGAC
     201                                                 250
KC   CAGTGATCAGTCCTCCCCAGAAATAGATGAAGACCGGATCCCCAACTCAC
IM   CAGTGATCAGTCCTCCCCAGAAATAGATGAAGACCGGATCCCCAACTCAC
     251                                                 300
KC   TTCTCAAGTCCACCTTGTCAATGTCTCCACGGCAACGGAAGAAGATGACA
IM   TTCTCAAGTCCACCTTGTCAATGTCTCCACGGCAACGGAAGAAGATGACA
     301                                                 350
KC   AGGACCACACCCACCATGAAAGAGCTCCAGACGAATGGTTGAACATCACC
IM   AGGACCACACCCACCATGAAAGAGCTCCAGACGA-TGGTTGAACATCACC
     351                                                 400
KC   TAGGGCAACAGAAGCAAGGGGAAGAACCTGAGGGAG-CACTTGAGAGCACA
IM   TAGGGCAACAGAAGCAAGGGGAAGAACCTGAGGGAGCCAC-TGAGAGCACA
     401                                                 450
KC   GGGAACCAGAATTCCTGCCACCCTGGGATTCCNAAAACAGGCTCGGGTTN
IM   GGGAACCAGGAGTCCTGCCCACCTGGGATGCCAGACACAGGCTCGGCGTC
     451                                                 500
KC   NAGGGCAAAATNCCCCNGGGAAANCNNAAANTTTTGCN-AAATCCCNCCCC
IM   AAGGCCCAGATACCCCGGGACAGCACAAAAGTCT-GCAGAA-TCCAACCCA
```

Figure A4. Comparison of the different sequences of clone 1912-d05. Sequences from King's College, London (KC), the partial sequence supplied by I.M.A.G.E. (IM), the MWG-Biotech (MW), and the DNASHEF sequence (SH) were compared. Differences are indicated by highlighted pink nucleotides: (N) is an unknown nucleotide (yellow); and (–) a gap for best alignment. There is degeneration of the sequences towards the 3' polyA tail.

```
     1                                                  50
KC   GTTCAGAGGCCAAGAATTCGGATCCATGGCCGCTGCCT-GACCGGGAGCCA
IM                  ATTCGGATCCATGGCCGCTGCCT-GACCGGGAGCCA
MW                                GCTGCCTAGACCGGGAGCCA

     51                                                 100
KC   TGGAGCCCGACAACAGCCCACGGAAGATCCAGTTTACGGTCCCGCTGCTG
IM   TGGAGCCCGACAACAGCCCACGGAAGATCCAGTTTACGGTCCCGCTGCTG
MW   TGGAGCCCGACAACAGCCCACGGAAGATCCAGTTTACGGTCCCGCTGCTG
SH   GTTAGCCCNACAACAGCCCACGGAAGATCCAGTTTACGGTCCCGCTGCTG

     101                                                150
KC   GAGCCTCACCTGGACCCGGAGGCGGCGGA-GCAGATTCGGAGGCGCCGCCC
IM   GAGCCTCACCTGGACCCGGAG-CGGCGGACG-AGATTCGGAGGCGCCGCCC
MW   GAGCCTCACCTGGACCCGGAGGCGGCGGAS--AGATTCGGAGGCGCCGCCC
SH   GAGCCTCACCTGNACCCGGATGTTGCGGAGC-AGATTCGGAGGCGCCGCCC

     151                                                200
KC   CACCCCTGCCACACTTG-TGCTGACCAGTGATCAGTCCTCCCCAGAAATAG
IM   -ACCCCTGCCACACTTG-TGCTGACCAGTGATCAGTCCTCCCCAGAAATAG
MW   CACCCCTGCCACACTTG-TGCTGACCAGTGATCAGTCCTCCCCAGAAATAG
SH   CACCCCTGCCACACTTGGTGCTGACCAGTGATCAGTCCTCCCCAGAAATAG

     201                                                250
KC   ATGAAGACCGGATCCCCAACTCACTTCTCAAGTCCACCTTGTCAATGTCT
IM   ATGAAGACCGGATCCCCAACTCACTTCTCAAGTCCACCTTGTCAATGTCT
MW   ATGAAGACCGGATCCCCAACTCACTTCTCAAGTCCACCTTGTCAATGTCT
SH   ATGAAGACCGGATCCCCAANTCACTTCTCAAGTNCACCTTGTCAATGTCT
```


251 300
KC CCACGGCAACGG-AAGAAGATGACAAGGACCACACCCACCATGAAAGAGCT
IM CCACGGCAACGG-AAGAAGATGACAAGGACCACACCCACCATGAAAGAGCT
MW CCACGGCAACGG-AAGAAGATGACAAGGACCACACCCACCATGAAAGAGCT
SH CCACGGNAA-GGGAAGAANATGACAAGGACCACNCCCACCATGAAAGAGCT
301 350
KC CCAGACGATGGTTGAACATCACCTAGGGCAACAGAA-GCAAGGGGAAAGAA
IM CCAGACGATGGTTGAACATCACCTAGGGCAACAGAA-GCAAGGGGAA-GAA
MW CCAGACGATGGTTGAACATCACCTAGGGCAACAGAA-GCAAGGGGAA-GAA
SH CCNTACGATGGTTGAACATCACCTAGGGCAANAGANAGCANGGGGAA-GAA
351 400
KC CCTGAGGGAGCCACTGAGAGCACAGGGAACCAGGAGTCCTGCCCACCTGG
IM CCTGAGGGAGCCACTGAGAGCACAGGGAACCAGGAGTCCTGCCCACCTGG
MW CCTGAGGGAGCCACTGAGAGCACAGGGAACCAGGAGTCCTGCCCACCTGG
SH CCTGAGGGANCCACTGAGAGCACAGGGAACCAGGAGTCCTGCCCACCTGG
401 450
KC GATTCCCAGACACAGGCTCNGCGTCCAAGGCCCAGATACCCCCNGGAACA
IM -AT-CCCAGACACAGGCTC-GCGT-CAAGG-CCAGATACCCCCGGGA-CA
MW GAT-CCCAGACACAGGCTCGGCGT-CAAGG-CCAGATACCCCCGGGA-CA
SH GAT-CCCAGACACAGGCTCNGNGTCAAGG--CCAGATACCCCCGGGA-NA
451 500
KC GCACAAA-GTTCTGCAGAA-TCCAACCCCA-GACTCCAGGA--CA-TGTT-GTGTTG
IM GCACAAAAGT-CTGCAGAA-TCCAACCCCAAGACTC-AGGAG-CAGTGTG-GTGT-G
MW GCACAAAAGT-CTGCAGAA-TCCAACCCCAAGACTC-AGGAG-CAGTGTG-GTGT-G
SH GCACAAAAGT-CTGCAGAAATCCAACCCCAAGACTC-ATGAGCCACTGTGCGTNT-G
501 550
KC GAGCCCAGA-CAGAG-ATTCTTCAGCCCCCATGCTACCATTG-ATTCCCA-GG-A
IM GAGCCCAGAACAGAGGATTCTTCAGCCCACATGCTACCACTGGATTCCCA-GG-A
MW GAGCCCAGAACAGAGGATTCTTCAGCCCACATGCTACCACTGGATTCCCA-GGGA
SH GAGCC-ACAACAGAAGATTCTTCNGCCNACATGCTACCACTGNAT-CCCATGN--

```
     551                                                          600
KC   CCA-GCTTGGTCTGACAAT---TTG-TATC-GGGGTTNNCCCN-GCNTNTT-GAATTC
IM   GCAAGC
MW   GCCAGCTTGGTCTGACAGAA-GTTGGTATCCGGGGATCACCACTGCAGTGTGG
SH   GCCANNT-GGTCTGACANANN-TTGGTATCCGGG-ATCANNACTGCNGTGNNGAAA
```

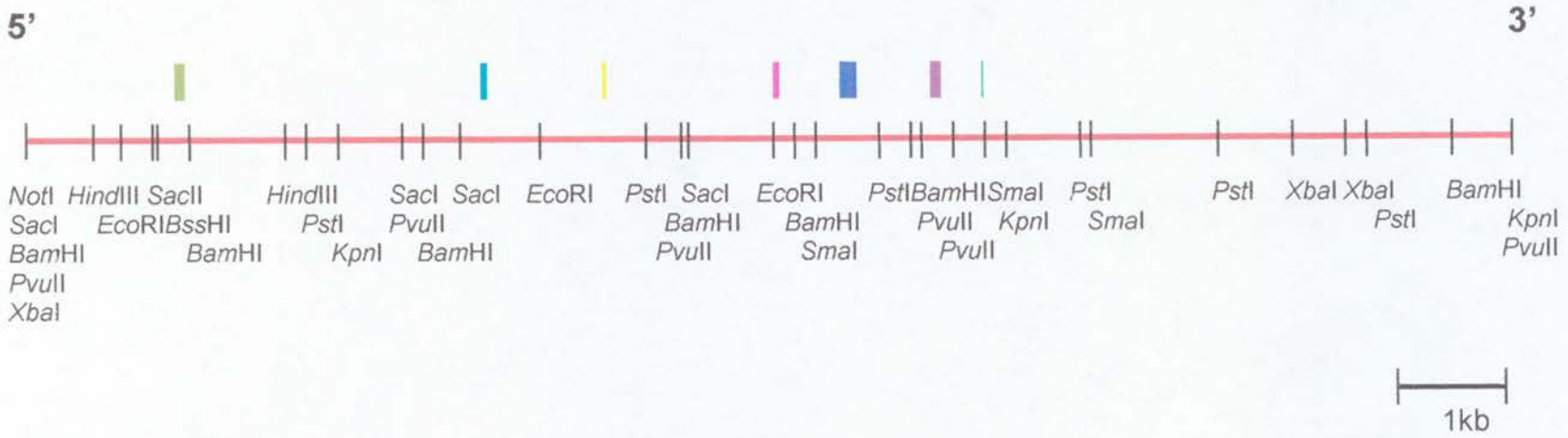


Figure A5. The restriction map of type 1 clone (14 kb, drawn to scale). The position of the exons (coloured boxes) and the scale bar are indicated.

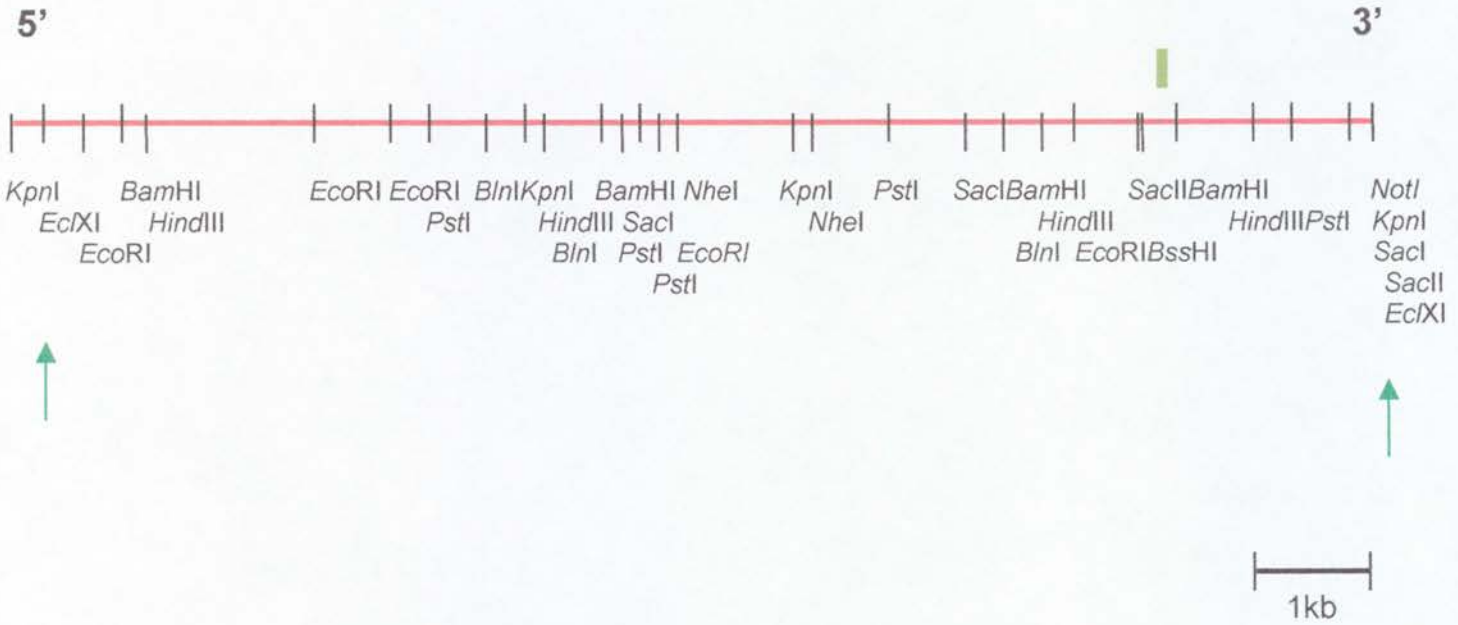


Figure A6. The restriction map of type 2 clone (12 kb, drawn to scale). The position of the exon 1 (coloured box) and the scale bar are indicated. The green arrows indicate the *Ecl*XI (5') and the *Not*I (3') restriction sites that were selected to insert the *Xho*I*Sal*I oligonucleotides.

Figure A7. Nine kilobases of sequenced genomic DNA. The genomic sequence of the region containing the I-1 protein-coding exons (1, 2, 3, 4, 5, 6 and 7) from the ATG start codon (nt position 1800) to the TGA stop codon (nt position 8709). The CpG island is underlined. The pBluescript vector sequences have been removed to allow the continuous sequence to be represented. N denotes a nucleotide of unknown type because of failure of sequencing. The *Sac*II site (CCGCGG) is at nt position 1511 (italics). A total of 9065 nucleotides were sequenced from subcloned fragments from the two clone types isolated from the library. Nucleotide 1 does not refer to the transcription start site since the promoter regions associated with the I-1 gene are not known, but does refer to the 5' flanking region of DNA that was sequenced.

```
1    AGGGGTGGGG ACTGTCCCTT CAGGTCATCC AGGCCTGGCC AAGAGGATAG
51   AGAATGTCCC TCTCTAGAAG CACTATGCCC CTAAGTGTAT GTGTGGGAGG
101  GAGAAGCAAA GAGAAGGAAA ATAATGGCCC AGCCCTAAGA CACCGTCTGC
151  CCTGTCTCCC ATATTATTGA TTCATCACCT CTCCGTGCAC ATTGCACTGG
201  GATGTGGGTA AACAGACATT GGGATCCAGG GACACTGAAC ATTTTCCCTA
251  CTATCCAAAC AGAGCAAAAA GGCTGGGAGA AGCATTCCCA GAGTCCCATT
301  ACTTCAACAG CAAATGGAGG AGCAACTGGC TATAGAACCT TCAGGGCTGC
351  TGAGTGGCTG CAACCTGGCC TTGCCCTGGA GAAAGCTGAG CCTGTCTGCT
401  GTGTAAAAGG ATCCCTAGGG GAGGAAGGGG GTGAGGGTGG GGCTATTTTT
451  GGAGGGAAGA GGGGAAGGGG CCATGGAAAA CACACCAGAG AACTCCAGAC
501  TGAGGTTTAA AGCTGCCCCC TAGAGTGGGG GATGACCAAG GCCCCATTCC
551  CCTGGAGCCA AGCTTAAACC TAGATAGACC CTGCCCTCTG TTCAGTTTAA
601  GACTACAAAG AAACTGATCT CCCCGTACTT CCATACCCCC CCTTCAGCAC
651  TATAGGCCTT CTCCTTCTCT ATCTCAGTTC CACTGTGCCT CCTGAATACA
701  TTAGGAACCC TTCCACCTCT CCAGCACCCA CCTCTCTCTG TCTAGTGACC
751  CTAACCCTAG TACAGAGCAC TTCAGTTCCT ATGAATATCC TGTCGCCTTC
801  ACTGGTTAGT ACCGAGTGGC ATGTTGATCC CATACCAACC CAGTAAGTTC
851  TTCTAGTAGT CTAACTTCCA TGCCAATAGT CACGACAGAG CAGGGCAGAG
901  CATTTTCATT GTCCTGCCAA CAGTGAAAGT GGACACCAGC CATCTGTAAA
951  CAGGCCACCC CGGTGGGGTG TCCTCCACTC CCCATCCCCA TCCCTGTAAT
1001 CCCCTTAAAG GCCTTTCACC TACTTTGCAG CTCACAATGT AGCAAGCAGA
1051 AAGTGAAATC TATTTCACAA GCAGGATTGG CTTTCACCTC CATGGCTTCT
1101 TCTTTACCCT AAATGGCCCC TCTCCCCAAG CCCTTGGTCC ACTCACACTC
1151 CAAAGAGTCT TACTCTGCCG ACTCTCTGAG AGCTTCCCCT CTTCTCTCTC
1201 TTGTCCAGGA GTGTTCCCTG AACAGTAGCC AAAGCTCGCA CCTGCCCTCA
1251 AAGTTTTCAC ATCCACTCTA GTAATTGGCC TCAAGGGGAA GTTCCCTCCC
1301 AGTCTAACTT TAATCCCTCC TGCTGTGTAC CTAAATCTCC TTTCAATCTT
1351 TTCTCTTTGT GAAATCTCCT TTAGATCATA AACATAATCT ACCTCTGGAG
1401 CTGTGTCCTT TCTTTTTTCC AGACCATGAA ACCTGCACCA GCTTGAAGAA
1451 GTTCTTGATA AACCTTTCCA CACCCACCCC TGCCCAGAAT CTAAAATACC
```

1501 TGGATGCCGG *CCGCGG*GTGG TCTAGGAAGC GACCAACCGG GGTCTGGATC
1551 CTTAGCTTGA GGGGCGGCAA CCTGCACAGC CGCCAGCACA GCCACGACGG
1601 GGTTAAGGCT CGGGACTGGG GGAGGACGGG GCCGGAGGAC GGTGACGAGC
1651 CAATCCGCAA CAACACAGTC CCCAAGCGGA GCGGGCGCGC CGAGCCAGGA
1701 GCTGCAAGCC GAGCGCCGGG CTGGGGCCAG GGCCGGAGCG CAACAGAGAG
1751 GGAGCGCGCC CGCCCCAGCC CGAGGTCCGC CGCTGCCTGA CCGGGAGCCA
1801 TGGAGCCCGA CAACAGCCCA CGGAAGATCC AGTTTACGGT CCCGCTGCTG
1851 GAGCCTCACC TGGACCCGGA GGCGGCGGAG CAGGTACAGG CAGGGCGGGG
1901 CAGCAGGACT CGGCGAAGGA CTGGGGGACC TGGCCGTGAG ACTTGAATCC
1951 CTGCCTTGCC CCAGCGCGGC ACCACACCTC TTCAGGAACG CACTCGCCCA
2001 GTTTATGGGG YTCTCCATAC CCTCCCCCAA CAGCGAGGCC ACACCTCGGT
2051 CACCAGTATG GGTCCCTGTG GCGGGTGTCT CCAAGAGCAT CGTTCACTCC
2101 CAGCCCAGCC ATAGGGACAC GACCCCTCCT CCCTCACTGG CGACCTTACA
2151 GCCCCTCACA CCCAGTCAAC CCCCAGAGTT TCCCGGCTAA GTGGCATTGG
2201 TAAGCTGAAG GGTACTGCCG GGAGGATGTT GCGGGCGGGG GGCGGCGAGT
2251 GAGCAGAAGA GACAGCTCTC CTCCCCTCCG TGTCATCTGG AGGCTGAGTG
2301 TGGCACCCAA AGGGATGAGA ACAGCTATGG AGGAGTTTTA GAGCAAGGAC
2351 GGGGGAGGGT GGAGGGCTCT GCAAGCTGAT GTTCCCTTCT GCAAGTCAGA
2401 GCACGTGGGA CCTGGGTATA ACTCTGCCCT ATGGAGTAGG GAGTGGGGGT
2451 GAGATGCGGG GCGCAGCGCG GCGATCTGGA GCGTCTGGGC CTGGGGTCCA
2501 AGCTTAGCCA CGAGCAAGCC ACCAAGAAAG CTTGGCTCAT CGGGTGGGAG
2551 TGCAGGGCCC TGGCCCAGCC TCCTGGCCTG GCCTATTTTT AGCTCTGAGC
2601 AGCCTGGCAG AGGCGGCAGT GTGTGAGTGA GCAGTGTGTG TCAGCGTCAG
2651 TGAGCGTATG GAAGGGGNNA CTATCTGCAG CCTGCCTTCT TTTGCATTCC
2701 CTGGATTGTG AAGACCATCC CTCCCCAAGA GCATCCAGCT CCCCAGGACT
2751 CCTGCCCCCA GCCACTCACT GTCAAGCTCT CCACCTGAAA GAATCAGACC
2801 TCTGGGGATG AGCAGGAGAA AAAGGTTTCC ACATTGCTAG AGCCGATTCA
2851 TAGAATTGGG TGTGGATGAA AAGGGACCCA GCAGGTGTAC ACTTTGAAAC
2901 TTCTACTTCT CAGATCCAGG GAGGAGCAGG NTACCCAGCT TCTCCTCTCC
2951 CTACCTGGGG ACCAAGGGGC TGGAGAGGGA GTTTGAGACC TCACACACAT
3001 GAGATGAGGA GAATGAAGAC TGTNGNTATG GACTATGCCT ATGCCCAGAA
3051 ACAGCAGATT GGCTCAGACG ACTTTTGAAA CAGGAAAGCA GGCCCAGAAG
3101 TGGAGACTGT GGGAGCCAGA GGTTCTGANA TCTGCCTCCC AATGGCAGAC
3151 AGGACCTGAC ACCTGCTTCT GTTTGGTGTC AAAGGGGTTT GAAGGGGCTG
3201 GTGAAATGGC TCAGTGGGTT AGAGCACCCG ACTGCTCTTC CGAAGGTCCA
3251 GAGTTCAAAT CCCAGCAACC ACATGGTGGC TCACAACCAT CCGTAACAAG
3301 ATCTGACTCC CTCTTCTGGA GTGTCTGAAG ACAGCTACAG TGTACTTACA
3351 TATAATAAAT AAATAAATCT AAAAAAAAAA AAAAAAAAGG GGGTTTGAAG
3401 GAGGGCTGGC CAGCTGGTTT ACAGAGCTGG GGACAAGGGA ACACAAGAGG
3451 GCACATGACA GTGTTGGGGA GTCAAAGTGA GAGAAAAGCT GGGACTCTGC
3501 TGCTCTTTTC TAGTCTTCAA AAATCTGCTC CTGCTCTTAA AGTTTGNCCC
3551 GGTTCTTCCT GCTACCCCAA CTTCCATGGG GTGGAAGTGG GAGGGTGAAG
3601 AAAAAGTGGG AAGCTAAGTG TACAGCANAG TGTGGGGACA GCCCTCCTCC
3651 ACAACCTTCA NATGTCCTTG AGAGCTAGGT TAGGTAGGGG CAGAAACAGA
3701 AAAGAGCCAT CTCTCTTGGG ATCCAGGGAG TCTCCTCCTG CCTCCTGCCT
3751 CCTGCCTCCT GCCTCCTGCC TCCTGCCTCC TGCCTCCTGC CTCCTGCCTC
3801 CTGCCTCAAC ACCACAATTC TCAAAGACTC AGAGTGCTTC AAATCCCTGT
3851 CTGGGGTTGG GGGTTGGGGG AGTCTGTTAA GTAAGGTAAG TAAGAGTCTG

3901 TCCAACCTCA GAGCCCTGCT GACCTTTGGT CATGGGCTCC TTGGGTTCCC
3951 AATACAGGTA CTGCATCAAG TCACTGATAC CCCTTNCACA GTCTCAGTCA
4001 CTTCCACTAC TCTCCAGTGG CAACCCCACC CCACCAAAAC AGAANGAAAG
4051 TTAGGGATCA AAGGGATGGT CCCACTTCCA CATACACGAN CATNTTGTAC
4101 CAGAAGCAGC CAGAGACAAT TCCATTCAAT TTTCACACCA CACACNCACA
4151 CNCAGANGTT NGCTTTTTGG GANCTTGAAT GCACTAAAAC NTATGCNCTA
4201 GTTTCCTGAN CAAAANTCAT TTNTCATTNA TCTGTTTTNT TTGGAAAANC
4251 CCTTTNGTTA AAANNCTTCT ACATTCNNTT TAGGNNCTAG GTATCTTGTC
4301 ACCTCCAGCC ACCACTGCCA ATGCTGATGA TGACAAGAGA CAAATATGNA
4351 CAGATCTAAA TCTGATCCGT TGCGTCCTCA CTGGACCCTG CTTCCCCAGG
4401 ACTGCTAGCT TTATCTACCC ACAGCCCTGA GAGAACAGCA GGAGCTCCAA
4451 AACGGCNNTT AAACCCATGT GATGCCTCTC TCTTCCAAAC CTTCCCCAGA
4501 AAAGGAAAAA CTACACATCC TAAGAGTCTA TAGGATTCAG CAACAATCAG
4551 CTTAATAGAT AGCAGGACAG GAAGAAAAGC CAGCAGCAGG AGGGGGGCAT
4601 TCAGTTGGAA CTCATGCCAC CGGTCTGCTT TATATGATCA TTTCCCCTGG
4651 GGCTGGCCAG GATTCTACCT CCTCAAGTCC TGTTCACCAT TAGGACCACT
4701 TCCCCTATTC CTGGCCAATT AAAAAGCCAT TGTGTCCAAG GCTAGTGAAA
4751 CCTTAGAGTA TTGAGGGACC GTGGTGGCCA GGGGAATGGT GCACTCAGCC
4801 CTTCCTTCAT GGTGGATGAA CCTGAGGGAG GCCTTCTGAA GACAGGTATG
4851 AGCCCACAAA GGGTTAATGT CTGCTTCAGG GGTGGACTGG AGTCTGAGGG
4901 AAAGCCCTCT GATCTCTCAT GTTATTTCTG ACCATTGATC CAGGATGGAC
4951 AGAAATCAAT GCACAGAGTG GACAGAGTTT CCGAAAGGAG AGCTGGGGAG
5001 GACAAGGAGA GGATGAAGAG ACAGTACAAG GAAGGGACTG TGTAGTTTCT
5051 TGTGTCCCCA AGAGCAGCTC ACCCCTGAAG AAGGGAATGG GCCAATGGAA
5101 CTCCAGAGTA ACCAGTAACT TTTGGGGGGA TGTGCAGAAT CGGACTGGGG
5151 GGGGGTATTC CTAGCGAACT TGACCTCTTA TTCTTCCTCC CCTAACCCAG
5201 ATTCCGCAGGC CGCGCCCCAG CCCTGCCACA CTTCTGCTGA CCAGTGATCA
5251 GTCCTCCCCA GGTAACCCCT TGTAAGAGGG GATGGGTGCT AGAGAGATAG
5301 GGCAGACCTC AGTAGTAGAG ATTTCCCCGT GCTGACTCAG CAAGATTTAT
5351 TTTGCTTTTA CAAACATGGA AATCTGCTAG GCGGGGAGGG GGACACCCTG
5401 ATTCCCTAAG CCCCCAGGCA TCAGGATGGT CAATTTAGTG CAGGTTGCTG
5451 TGACAACACC CTTGTAGTTT TAGTGCAGGT TGCTGTGACA ACAGCCTTGT
5501 AGTTTTCTCC AGCAACTAAG ATGTCTGCTA TTTCCACTCC AAATGCAGTA
5551 GTGACAGCTG TGAGGGGGAG AGATTGTCTC CATCATCTAG CATCAGGTCC
5601 CCAGAGGGTG AGGTTAGAAA CAGCTCAGGG AGCAGCTAGA ATTCTGGTCA
5651 CCAGACTCCC CACCAGGGTC TGAGGACACA AGGTCTTAGA GAATGTCAGA
5701 GGAGCTTGGT AAGTGTAAAA AAGTGACCAG GCTAGTCTGA AATCGGGCCA
5751 AAGTCTCCAG CTCTCCCCAC CCCAACTTTC ATAAATGGTG CACATCTATG
5801 GACCCTGACC AGAATGAGTT TGCTATTCTC TCTCAGGGAG CCCTGCTCCT
5851 GTGAAGGCTC AGCTCTGGTC TGGGGGTAAA AGAATGACTA GAAGGGAAAA
5901 GAGACCTGGA TGCTGCAGAG GAGGAAACAA AAACAAAAAC AAAAACGCCG
5951 AAGGGCCACA CATTTGCCTT CACAGTCTTC AGAGCCTGAG GCCAGGATGC
6001 AGGCTTAGGC TGGTCAAAGG GTCCACTCTG TTTTGACAAA GCCAAGCCAC
6051 AGTCAAGGGG GATCTGCTGT GTACCTTCTT CTTAAGAACT CGGTCTTTGA
6101 NGAAACTTTT GGGTTTCCTG TCTTATCACT GGTTTTCCCT ATTNCCCCAG
6151 AAATAGATGA AGACCGGATC CCCAACTCAC TTCTCAAGGT GAGCTGGCCC
6201 CTACCATTCA GCCCAGGCCT CAAAAGCCAA ACCCTCATCC TCCAGATCTG
6251 GGGTGTAGCT CGGTAAAGAG AGCACCACTT AGCATATACA GACCCAGAGC

6301 ACCCATGAAC CCAAAACAAA AGCCTTGTAC CTCACCAGTA GGACCTTTGA
6351 TGTCCTCTTC TATGCACCCC AAAAAATGCA TGCCCCCCTG CCATCATACA
6401 CTGTCTCCAA GCTCACCCTG TTGGAAACAA GAATTTCCTC TCGCTCCTAA
6451 AACCCACCTG CCCTAATCTC TAATCCAAAA TCATGCCAAA GCAGAAAAGG
6501 GTGGGAAGAG CTGGCACCCA TCCAGACTCT GTGTCCTTCA CAGTCCACCT
6551 TGTCAATGTC TCCACGGCAA CGGAAGAAGA TGACAAGGAC CACACCCACC
6601 ATGAAAGGTT TGGCATCCCC TTCCCAATTT CTCATCTGGC TGAAAGTGGA
6651 GAGTAGATCT AGTAACCCTT TGCAGGGAGT CAATGGGAGG TTACTGGGTC
6701 TGCCTTTATT TGAGTGGAGT TATTGCGATG TGAAGAAGGC TCCTGTAGCT
6751 TGAGGAAGCT GGTCCCTAGA CCAAACCAGA ATACCACCTC CTCACTATGA
6801 CAGAGAAATA GTACATGTAA CCTCATCAGC ACAGCTCATG CAAGGTACTA
6851 TATTCCCTGA GTCCAAATTG AGGCTTCTCT CCTGCCTCTA AGGAGAGACC
6901 ACAGGTCCTA GTGCATTAGG AGTGTCTTCT GCACTGAAAG GCTCTTCCCC
6951 TCTGTTTGCC TTTGAAAAGA CATTATGCTC CTGCTTTCAT CTCTCCTTAT
7001 TGCCCAGCAA ATGTTCCTCC TAATTTAGAG CCCAGTTCTG GGAAATTAAA
7051 GCACCCAGCA ACATGCACAG ACCCAATGGT GGGACTAGTG AATTTAGGAG
7101 ACCTCTCACA GGAGATTTTA GGCAGTTATA TCCAATGTTT CCTCCCTTTC
7151 TCTCTGCCTC CAGGGCGATT GGAGCTCCAG ACGATGGTTG AACATCACCT
7201 AGGGCAACAC AAGCAAGGCC AAGAACCTGA GGGACCCACT GAGAGCACAG
7251 GGAACCAGGA GTCCTGCCCA CCTGGGATCC CAGACACAGC CTCGGCGTCA
7301 AGGCCAGATA CCCCCGGGAC AGCACAAAAG GATGTACAAT CCTGGGAGGG
7351 TAACAGAATG CAGTTCCCCA GCCTTGCAGG AAGACAGCTT TACCCACTGG
7401 GGTATAGGGA GGAATCACAG CCTGAGAGGG TGGAATAATA GCTAAAAGCC
7451 CAGTCAAGGA CATGGGTGTA TTCCATACAG CATGATTTGA CTCACTTTGT
7501 CTAATTGTAG TTACCACCTT AGACCCCAGA GATCCTGCCC TGGCTGGAAA
7551 AGGAGGGAAA CCCCATCAGA CTTAGATCAG TACTTGACCC TGTATGGGAG
7601 CCAGGGCTAC AGTTCTGCCT TCACCGCTGG CCACTTTACC AGCAGGGCTG
7651 GCGTCTATTT TGCTGGAGTT GTTTTGGAGT CTTCCAATGA GTGTGGCTCA
7701 ACACCTGACA CATACAAAGC CTGCCATACA AGGAATGTGG TGCTGGTGGT
7751 GCTAATCCAC AGCCACATTC CCTTCTCCCA GATGCGTTCC AGAATTCAGA
7801 CCATTTCCAA TTGTGTGAAA GTGATATAGT ACACACATTT CCTGTTACAT
7851 AACACACCCA GCAGAGCATC CTGTAATTAA GCACACTCTT TTCCTACAAT
7901 CCAGTGTCTA ACTGGGATAG GTAACATCAC TTGAGTTTAA ACTCTAATTT
7951 TTTTTTATTT ATTTACATAT TTAGTTAGTG AGTTTTTAGG GTTTGGGGGG
8001 TTCGTTTTGT TTTATTTTTA CTTTTTAATC TCTGGTTACA GGGGGGAAAA
8051 TAGTGAAGCT ATCATAATAA CTGTGATCAG AAAAACCCCA GACGGCAAGG
8101 TAGATCTGCT CCCAATGGGT CTTGGGCCCA TTCTCAGCAG CTGCACACCT
8151 TTATAAGGGT TGCTCTCCCT CCATGGCCCA CTTTCCCTCC CTAGAGTCTG
8201 CAGAATCCAA CCCCAAGACT CAGGAGCAGT GTGGTGTGGA GCCCAGAACA
8251 GAGGATTCTT CAGCCCACAT GCTACCACTG GATTCCCAGG GAGCCAGCTT
8301 GGTAAGCCCC TTGGCCTAGA CTTTGGGGCC TTCAGGAGAG CAGGCTGTAT
8351 GGGTTAGTAT TTGAAGCAAA CGTGTAGACA AAGAAGCAGC TTTATTTATT
8401 CACTAAAAAG TTAGCCTTTG TGCTAAGAGA CTCCCTCTGA CAAAGAGAGG
8451 AGGGATTAGG TGACTGCGGA GAGGATCAAA CGGAGAGTTT GTTGTGTAGA
8501 CCTTCTGGAA CAGGTGATGG AAATAGAGCA TAGACTTAAT TTTATTTTTT
8551 AATAAGGGGC AGTTGGAACC AAAGTACTGT GGGCCTAGAA GAGATGGGGC
8601 TCGGACTGAG ATCCACAGCA GGACAGAAGG TGAATAGCTG GGCTGGCCAG
8651 GAAGAAGGAT GAATCTATGG CCCTCAGTCA CCTCCCCTTT CCTCCCCCAA

```
8701 AACAGGTCTG ACAGAAGTTG GTATCCGGGG ATCACCACTG CAGTGTGGAA
8751 ATTCATGGAC ACTGGATGTT TCTTAATCTC TTGTTTTTAA AAAGTGATAA
8801 ATTTAGGTGT TAGGTCCTTG GAGCTTTTCT TCTTTCCAGC CCGGGGATCC
8851 CTTCCTCCTT CAGCTCTTGT CTGCACCCCA CTGCCCCCAC CCTCAGGACC
8901 GATTAAGGAA GGAAATAGCA AACTGACAAA AGCCTAGCCT CCACTAGCAA
8951 AATAGAAAAA TAGACGTCTG GAAGTAGGAG GTCGCCTCAA GGTTTTCCTT
9001 TGNNCTGTGG AAAGCTCTGC CCACAGAGCA GGTTAGCCTA CTTGGGAGAC
9051 TGCTGCTCTA AAGCG
```

Figure A8. Comparison of the 2kb *Hind*III restriction fragment from clone type 1 (T1) and clone type 2 (T2) isolated from the library screening. Since the two sequences are almost identical, the two clone types overlap.

```
      1                                                  50
T1    GGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGGACCCCAGGCCCA
T2    GGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGGACCCCAGGCCCA
      51                                                 100
T1    GACGCTCCAGATCGCCGCGCTGCGCCCCGCATCTCACCCCCACTCCCTAC
T2    GACGCTCCAGATCGCCGCGCTGCGCCCCGCATCTCACCCCCACTCCCTAC
      101                                                150
T1    TCCATAGGGCAGAGTTATACCCAGGTCCCACGTGCTCTGACTTGCAGAAG
T2    TCCATAGGGCAGAGTTATACCCAGGTCCCACGTGCTCTGACTTGCAGAAG
      151                                                200
T1    GGAACATCAGCTTGCAGANCCCTCCACCCTCCCCCGTCCTTGCTCTAAAA
T2    GGAACATCAGCTTGCAGAVCCCTCCACCCTCCCCCGTCCTTGCTCTAAAA
      201                                                250
T1    CTCCTCCATAGCTGTTCTCATCCCTTTGGGTGCCACACTCAGCCTCCAGA
T2    CTCCTCCATAGCTGTTCTCATCCCTTTGGGTGCCACACTCAGCCTCCAGA
      251                                                300
T1    TGACACGGAGGGGAGGAGAGCTGTCTCTTCTGCTCACTCGCCGCCCCCCG
T2    TGACACGGAGGGGAGGAGAGCTGTCTCTTCTGCTCACTCGCCGCCCCCCG
      301                                                350
T1    CCCGCAACATCCTCCCGGCAGTACCCTTCAGCTTACCAATGCCACTTAGC
T2    CCCGCAACATCCTCCCGGCAGTACCCTTCAGCTTACCAATGCCACTTAGC
      351                                                400
T1    CGGGAAACTCTGGGGGTTGACTGGGTGTGAGGGGCTGTAAGGTCGCCAGT
T2    CGGGAAACTCTGGGGGTTGACTGGGTGTGAGGGGCTGTAAGGTCGCCAGT
      401                                                450
T1    GAGGGAGGAGGGGTCGTGTCCCTATGGCTGGGCTGGGAGTGAACGATGCT
T2    GAGGGAGGAGGGGTCGTGTCCCTATGGCTGGGCTGGGAGTGAACGATGCT
      451                                                500
T1    CTTGGAGACACCCGCCACAGGGACCCATACTGGTGACCGAGGTGTGGCCT
T2    CTTGGAGACACCCGCCACAGGGACCCATACTGGTGACCGAGGTGTGGCCT
```

```
      501                                                 550
T1    CGCTGTTGGGGGAGGGTATGGAGAYCCCCATAAACTGGGCGAGTGCGTTC
T2    CGCTGTTGGGGGAGGGTATGGAGAYCCCCATAAACTGGGCGAGTGCGTTC
      551                                                 600
T1    CTGAAGAGGTGTGGTGCCGCGCTGGGGCAAGGCAGGGATTCAAGTCTCAC
T2    CTGAAGAGGTGTGGTGCCGCGCTGGGGCAAGGCAGGGATTCAAGTCTCAC
      601                                                 650
T1    GGCCAGGTCCCC-AGTCCTTCGCCGAGTCCTGCTGCCCCGCCCTGCCTGTA
T2    GGCCAGGTCCCCCAGTCCTTCGCCGAGTCCTGCTGCCCCGCCCTGCCTGTA
      651                                                 700
T1    CCTGCTCCGCCGCCTCCGGGTCCAGGGTGAGGCTCCAGCAGCGGGACCGT
T2    CCTGCTCCGCCGCCTCCGGGTCCAGG-TGAGGCTCCAGCAGCGGGACCGT
      701                                                 750
T1    AAACTGGATCTTCCGTGGGCTGTTGTCGGGCTCCATGGGCTCCCGGGTCA
T2    AAACTGGATCTTCCGTGGGCTGTTGTCGGGCTCCATGG-CTCCCGG-TCA
      751                                                 800
T1    GGGCAGCGGGCGGGACCTCGGGCTGGGGCGGGGCGCGCTCCCCTCTTCTG
T2    GG-CAGCGG-CGG-ACCTCGGGCTGGGGCGGG-CGCGCTCCC-TCT-CTG
      801                                                 850
T1    TTTGCGCTCCGGGCCCTGGGCCCCAGCCCGGGCGCTCGGCCTTGCAGCTC
T2    TT-GCGCTCCGG-CCCTGG-CCCCAGCCCGG-CGCTCGGC-TTGCAGCTC
      851                                                 900
T1    CTGGCTCGGCGCGCCCCCGCTCCGCCTTGGGGGAGCTGTTGTTGTTGCGG
T2    CTGGCTCGGCGCGCCC--GCTCCGC-TTGGGG-A-CTGT-GTTGTTGCGG
      901                                                 950
T1    ATTTGGCTCGTTAACGGTCCCTCCCGGCACCCTTCCTCCCCCAAGTCCCG
T2    ATT-GGCTCGTCA-CCGTCC-TCC-GGCCCCGT-CCTCCCCCA-GTCCCG
      951                                                1000
T1    AA-CCTTAACCC-GTCGTGGCTGTTCTGGGCGGCTCTTGAAGGTCGCCGCCC
T2    -AGCCTTAACCCCGTCGTGGCTGTGCTGG-CGGCTGT-GCAGGTTGCCGCCC
```

Appendix 2

LSEVIER

Mechanisms of Development 96 (2000) 237–241

www.elsevier.com/locate/modo

Gene expression pattern

Xpression and genomic characterization of protein phosphatase inhibitor-1: a novel marker for mesothelium in the mouse

Lorna McLaren[a,*], Shelagh Boyle[b], John O. Mason[a], Jonathan B.L. Bard[a]

Department of Biomedical Sciences and Centre for Developmental Biology, Hugh Robson Building, University of Edinburgh, Edinburgh EH8 9XD, UK
[b]*MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK*

Received 2 May 2000; received in revised form 13 June 2000; accepted 13 June 2000

ostract

Protein phosphatase inhibitor-1 (inhibitor-1 or I-1) is involved in signal transduction and is an endogenous inhibitor of protein phospha- e-1. The mouse I-1 protein sequence has been deduced from cDNA and is strongly homologous to the published rat sequence. A mouse nomic library was screened, and the I-1 gene was characterized and localized by fluorescent in situ hybridization (FISH) to chromosome F. Protein expression in a range of embryonic and adult tissue was analysed using confocal microscopy. Inhibitor-1 is expressed by: the elomic epithelium; the epithelial bounding layer of cells of the kidney, lung, liver, heart, intestine and gonad; and the surface ectoderm. e blast cells of the kidney do not express I-1. We conclude that I-1 is a marker for mesothelium. © 2000 Elsevier Science Ireland Ltd. All hts reserved.

ywords: Protein phosphatase inhibitor-1; Inhibitor-1; I-1; Chromosome; mRNA; Gene; Mesothelium; Ectoderm; Laminin-1

Results

1. Genetic analysis

A keyword search of the dbEST database yielded a full- ngth mouse cDNA clone homologous to rat protein phos- atase inhibitor-1 (inhibitor-1 or I-1; dbEST clone 926702; nnon et al., 1996). Sequencing this clone allowed us to duce the 171-amino-acid sequence of the protein, which ows 95.3% homology to the rat sequence (Fig. 1). Using ST sequences as a probe, two types of genomic clones ere isolated from a library screening: the type 1 clone 4 kb) spans the entire I-1 coding sequence; and the type clone (12 kb) includes the first exon and ~9 kb of the 5′ nking sequence (Fig. 2). Subclones from type 1 DNA ere sequenced and analysed to identify exon location. nalysis of ~400 bp of the 5′ sequence (including the st exon) using the NIX computer program (http:// ww.hgmp.mrc.ac.uk/) revealed a CpG score of 0.74. The gorithm is based on work by Gardiner-Garden and From- er (1987) using a scale between 0 and 1.0, with values oser to 1.0 indicating the presence of a CpG island. Fluor- cent in situ hybridization (FISH) showed that the I-1 gene (Fig. 3a) is located at the end of chromosome 15, band F (Fig. 3b).

1.2. Protein expression

Svennilson et al. (1995) suggested that I-1 is a specific marker for kidney mesenchymal stem cells. Therefore, the I-1 expression pattern in mouse kidneys was investigated (Table 1). From E12.5 onwards, the protein is restricted to the cytoplasm of the peripheral layer of cells (Fig. 4a,b) that has a unilaminar squamous epithelial morphology (Williams et al., 1995). However, in E11 and 11.5 kidneys, I-1 is expressed by the coelomic epithelium surrounding the intermediate mesoderm within which is the early kidney rudiment (Fig. 4c).

The Pax-2 antibody labels nuclei of condensed metanephric mesenchyme at the kidney periphery (Dressler and Douglass, 1992). E13.5–15.5 kidneys were analysed for Pax-2 expression (Fig. 4d), but since the antibody did not label the outermost layer of cells, there is no co-expression with I-1. Thus, I-1 does not label metanephric mesenchyme cells.

The epithelial morphology of the I-1-positive cells was confirmed with anti-laminin and anti-pan cytokeratin antibodies using E12.5–16.5 kidneys. The laminin antibody labelled the basal lamina of the peripheral I-1-positive cells and developing nephrogenic tubules (Fig. 4e,f).

Other organs from the coelomic cavity (e.g. lung and

Corresponding author. Tel.: +44-131-650-3106; fax: +44-131-650-45.

E-mail address: lornam@holyrood.ed.ac.uk (L. McLaren).

: S0925-4773(00)00388-9

```
                1                                                   60
MOUSE   MEPDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEIDEDRIPNSLL
RAT     MEPDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEVDEDRIPNPLL
HUMAN   MEQDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEIDEDRIPNPHL
RABBIT  MEQDNSPRKIQFTVPLLEPHLDPEAAEQIRRRRPTPATLVLTSDQSSPEVDEDRIPNPLL
        ** ************************************************ *******  *

               61                                                  120
MOUSE   KSTLSMSPRQRKKMTRTTPTMKELQTMVEHHLGQQKQGEEPEGATESTGNQESCPPGIPD
RAT     KSTLSMSPRQRKKMTRTTPTMKELQTMVEHHLGQQKQGEEPEGATESTGNQESCPPGIPD
HUMAN   KSTLAMSPRQRKKMTRITPTMKELQMMVEHHLGQQQQGEEPEGAAESTGTQESRPPGIPD
RABBIT  KPSLAMSPRQRKKMTRTTPTMKELQMMVEHHLGQQEQGEEPEGAAEGTGAQESQPPGTPG
        *  * ********** ******** ********* ******** * ** *** *** * *

               121                                                 171
MOUSE   TGSASRPDTPGTAQKSAESNPKTQEQQGVEPRTEDSSAHMLPLDSQGASLV
RAT     TGSASRPDTSGTAQKPAESKPKTQEQRGVEPSTEDLSAHMLPLDSQGASLV
HUMAN   TEVESRLGTSGTAKKTAECIPKTHERGSKEPSTKEPSTHIPPLDSKGANSV
RABBIT  TGAESRLGPSATAQKPAQPSPRAQERRGEEPSTAKTS-Q----DSQGASAV
        *   **     ** * *   *   *   ** *   *      ** **  *
```

Fig. 1. Comparison of mouse inhibitor-1 amino acid sequence with the predicted sequence from rat skeletal muscle cDNA (Elbrecht et al., 1990), the deduced human sequence from cDNA clones isolated from a brain library (Endo et al., 1996) and the sequenced protein derived from rabbit skeletal muscle (Aitken et al., 1982). The alignments were obtained by Clustal analysis (http://www.hgmp.mrc.ac.uk/). Asterisks indicate residues which are identical in all four species. The 25 nucleotide differences between the mouse and rat mRNA sequences result in eight amino acid differences (indicated by shaded text) only one of which is conservative (amino acid 50).

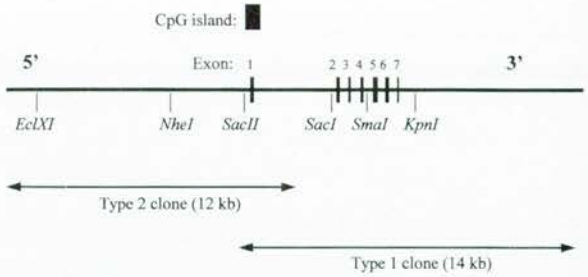


Fig. 2. Two clone types were isolated from the λ129 mouse genomic libr and span the inhibitor-1 (I-1) gene with a 2-kb overlap. The CpG island a selected restriction sites are marked. The I-1 gene is contained in ~7 kb genomic DNA and the deduced protein sequence is encoded by sev exons.

heart from E12.5 to adult) express a peripheral laminin-positive basal lamina (Fig. 5a) that underlies the I-1-positive cells (Table 1 and Fig. 5b–f).

The developing gonads express I-1 in the coelomic epithelial layer from E12.5 (Fig. 6a,b). In the testis, I-1 is also expressed by the sex cords, but is down-regulated after their formation: at ~E13.5, three out of 12 testes were positive for I-1 expression in the internal sex cords (Fig. 6c–e), but after E13.5, there was no further internal I-1 expression.

Inhibitor-1 is also expressed in the outermost layer of cells of the ectoderm and the periderm of the cornea, an ectoderm-derived layer (Fig. 7 and Table 1). However, I-1 expression in brain and neural tube (the other major derivatives of ectoderm) was not visible in whole-mount embry (E9.5–11.5), nor in embryonic mesenchymal derivativ (e.g. bone and tendon), vascular endothelium or epitheli tissue such as lung bronchii (data not shown).

2. Conclusion

The localization and immunostaining observatio suggest that I-1 is expressed by the mesothelium, a unil minar squamous epithelial cell layer derived from the coel mic epithelium that covers the tissue of the coelomic cavi

3. Experimental procedures

3.1. Genetic analysis

The dbEST database was screened to obtain mou cDNA clones homologous to the rat I-1 sequence. Clo

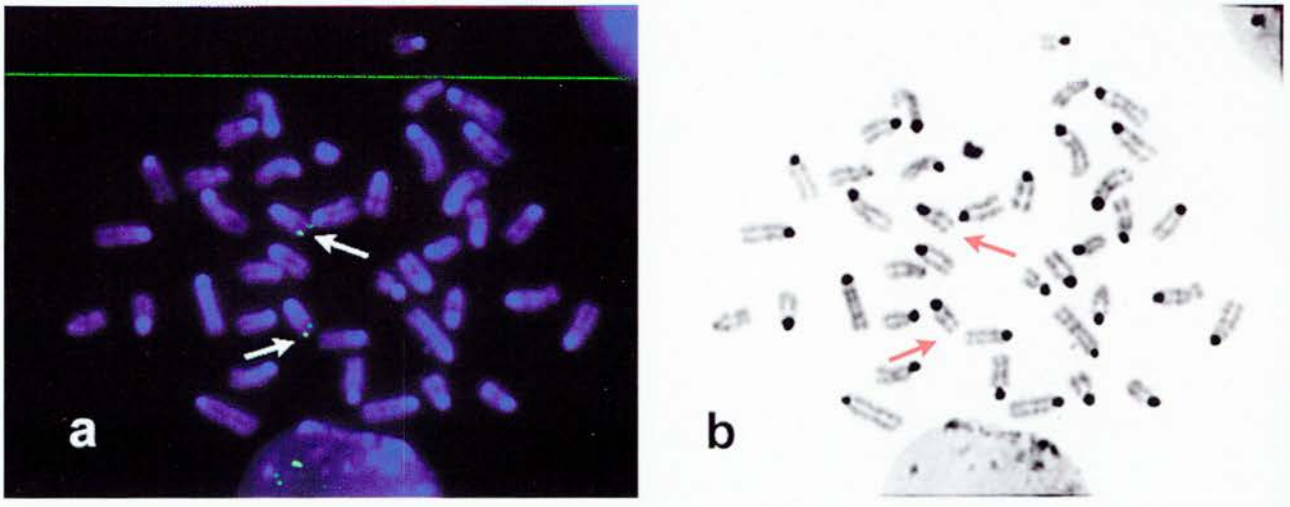


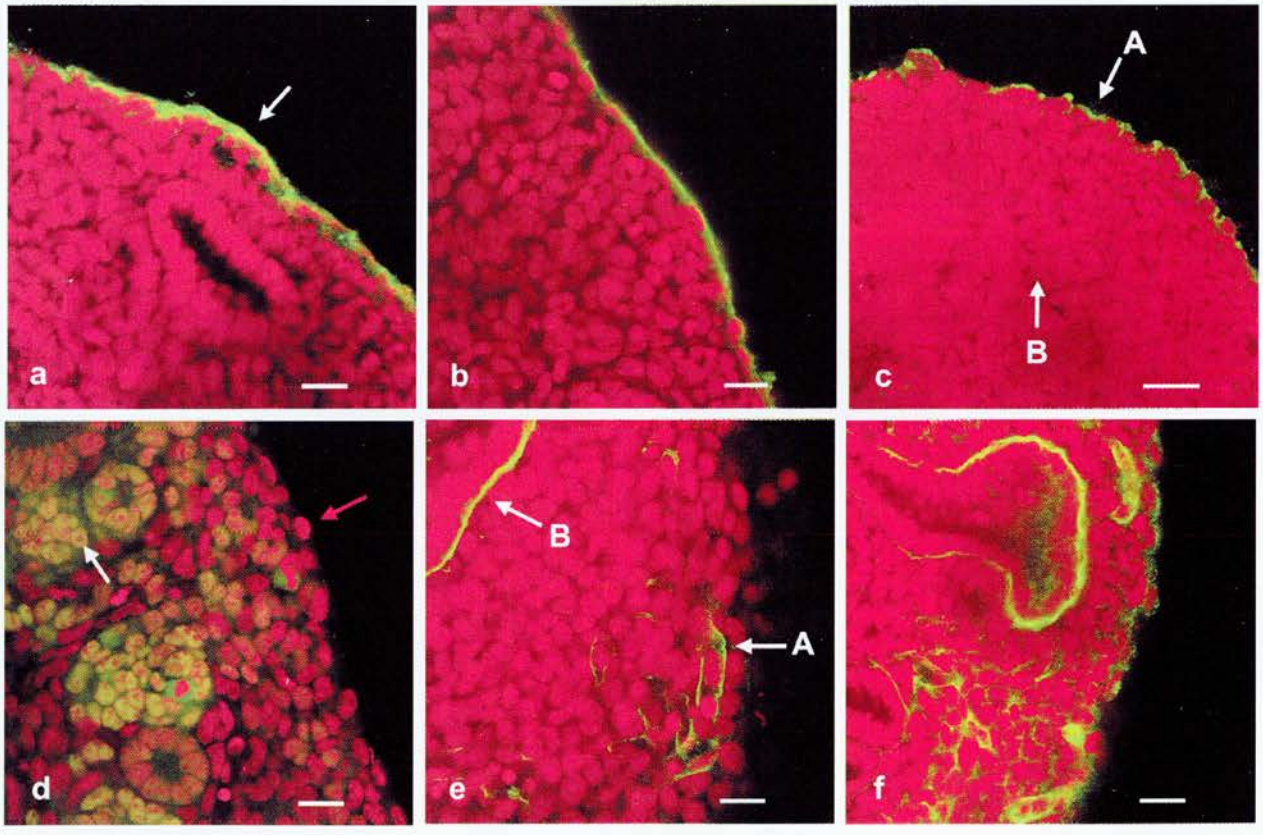
Fig. 3. Chromosomal localization of the inhibitor-1 (I-1) gene. (a) The genomic clone type 1 was labelled with biotin, and the signal visualized using success layers of avidin FITC, biotinylated anti-avidin and avidin FITC. The chromosomes were counterstained with 1 g l^{-1} 4′,6-diamidino-2-phenylindole (DAI and hybridization signals were visualized with a Zeiss Axioplan epifluorescence microscope. The white arrows show the I-1 gene on a normal mo metaphase spread and the I-1 signal is apparent in an adjacent interphase cell. (b) Giemsa banding of the metaphase spread showing the localization of I-1 gene on chromosome 15 band F (indicated by red arrows).

able 1
nalysis of mouse tissue for inhibitor-1 (I-1) expression

issue	Embryonic age of mouse tissue (days)[a]										
	9.5[b]	10.5[b]	11.5[b]	12.5	13.5	14.5	15.5	16.5	17.5	18.5	Adult
idney	ND	ND	+	+	+	+	+	+	+	+	+
ung	ND	ND	ND	+	+	+	+	+	+	+	+
iver	ND	ND	ND	+	+	+	+	+	+	+	+
eart	ND	ND	ND	+	+	+	+	+	+	+	+
ıtestine	ND	ND	ND	+	ND	ND	+	+	ND	+	+
onad	ND	ND	ND	+	+	+	+	+	+	+	+
kin	+	+	+	+	+	+	+	+	+	+	+
ye	+	+	+	ND	ND	ND	ND	+	ND	ND	ND

[a] +, positive staining for I-1; ND, not determined.
[b] Whole-mount embryo used.

26702 was sequenced and a 5′ 336-bp PCR product was mplified. A λ129 PS mouse genomic library (NBL Gene ciences Ltd) was screened using the 336-bp PCR product nd full-length clone 926702. DNA from the type 1 clone 'as subcloned, sequenced and analysed using the NIX computer program (http://www.hgmp.mrc.ac.uk/). The genomic DNA clone type 1 was labelled by nick translation with Bio-16-dUTP (Roche) and hybridized to normal mouse metaphase chromosomes, as described by Fantes et al. (1992).



g. 4. Confocal analysis of a range of kidney ages using antibodies to inhibitor-1 (I-1), Pax-2 and laminin (FITC, green; propidium iodide, red). Inhibitor-1 ;pression is shown in: (a) E12.5 kidney (bar = 20 μm); the I-1 signal is indicated by the white arrow and is restricted to the periphery of the developing organ; ›) E16.5 kidney (bar = 20 μm); and (c) E11.5 kidney (bar = 20 μm); I-1 expression is present in the epithelium of the coelomic cavity (white arrow, A), but ɔt in the rudimentary kidney (white arrow, B). (d) Pax-2 expression in the E13.5 kidney (bar = 20 μm). The nuclei of the peripheral condensed mesenchyme ells and collecting tubes are positive (indicated by the white arrow). The red arrow shows the peripheral layer of cells that does not express Pax-2. Laminin ;pression in: (e) E12.5 kidney (bar = 20 μm); there is weak laminin expression in the basal lamina of the peripheral cells of the tissue (white arrow, A) in ɔmparison to that around the ureteric bud (white arrow, B); and (f) E13.5 kidney (bar = 20 μm); peripheral laminin staining increases as the kidney matures.

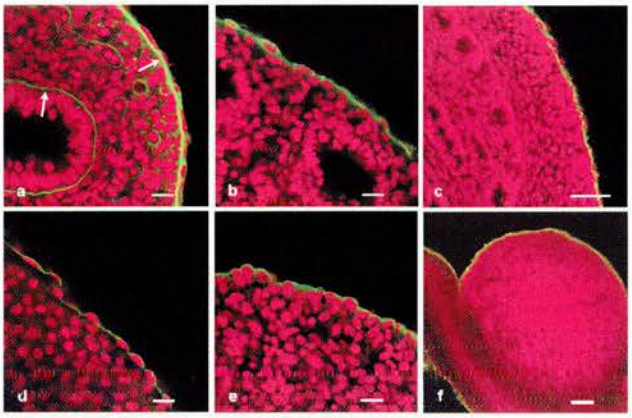


Fig. 5. Laminin and inhibitor-1 expression in tissue of the coelomic cavity (FITC, green; propidium iodide, red). (a) Laminin expression in E12.5 lung (bar = 20 μm). There is laminin expression in the basal laminae of the peripheral layer of cells and within the lung (indicated by white arrows). (b–f) Inhibitor-1 expression in embryonic mouse tissue: (b) E13.5 lung (bar = 20 μm); (c) E17.5 intestine (bar = 50 μm); (d) E12.5 heart (bar = 20 μm); (e) E12.5 liver (bar = 20 μm); and (f) E17.5 ovary (bar = 100 μm).

3.2. Protein expression

Noon on the day of mating is referred to as E0.5. Mouse

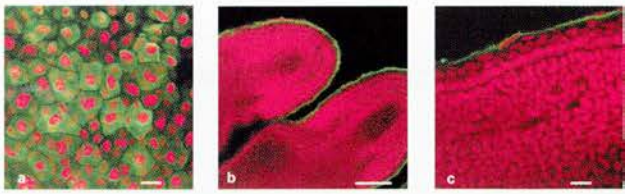


Fig. 7. Inhibitor-1 expression at E16.5 (FITC, green; propidium iodide, red showing (a) the surface view of periderm of the eye (bar = 20 μm) an (b,c) the surface ectoderm of hindlimb digits (b, bar = 100 μm; c, bar = 2 μm). Inhibitor-1 expression is restricted to the surface epithelial cells.

embryonic and adult tissues were fixed in 4% paraforma dehyde. The I-1 antibody (Serotec) was diluted in PBS/1% Triton X-100 (Sigma) and whole-mount specimens wer incubated overnight. The I-1 antibody was detected b anti-sheep FITC antibody (Sigma), and the tissue wa RNase-treated, counterstained with propidium iodide an analysed by confocal microscopy (Leitz TSCNT). Tissu incubated in PBS instead of PBSTX gave an identica expression pattern. For negative controls, tissues wer processed as described above, but no primary antibod was added. Anti-Pax-2 antibody (a gift from Dr Greg Dres ler) and anti-laminin antibody (Sigma) were detected b anti-rabbit-FITC secondary antibody (Sigma).

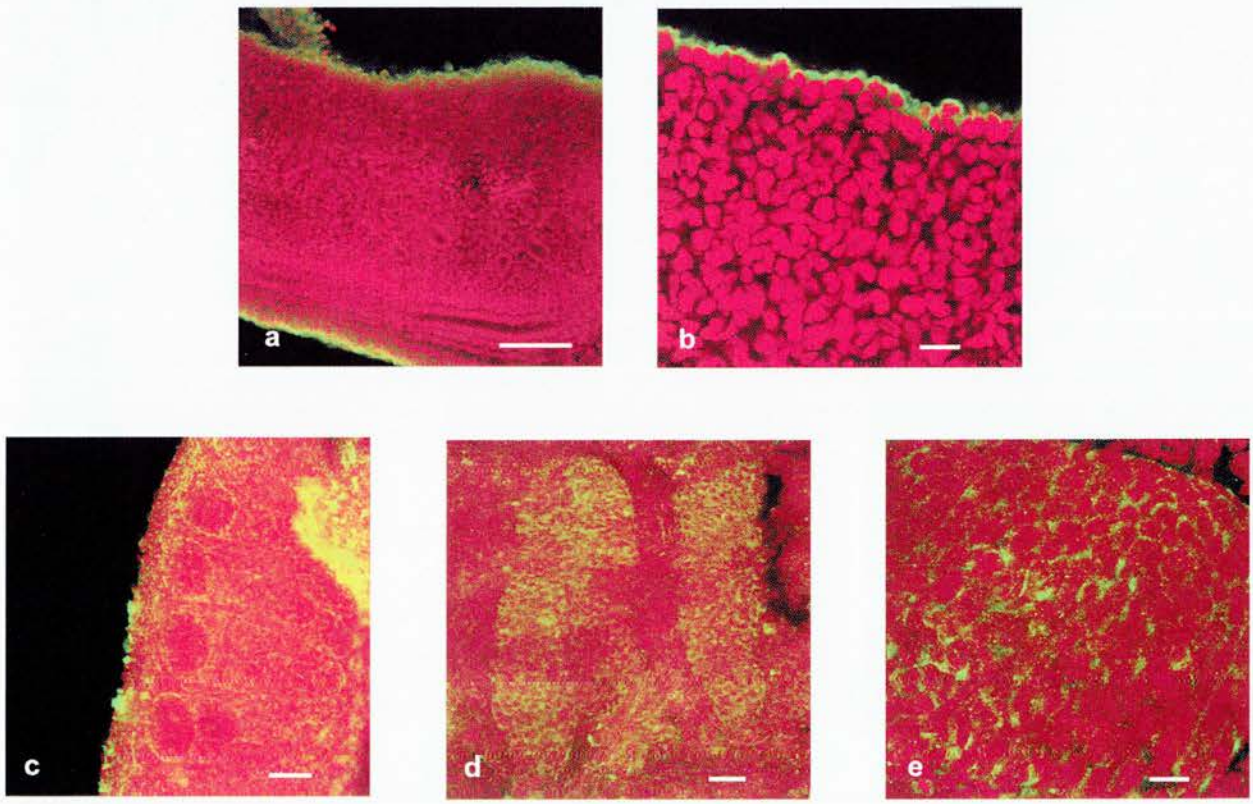


Fig. 6. Inhibitor-1 (I-1) expression in the developing gonads (FITC, green; propidium iodide, red). (a,b) Inhibitor-1 expression in E12.5 in different gonads restricted to the surface epithelium (a, bar = 100 μm; b, bar = 20 μm). (c–e) Inhibitor-1 expression in the sex cords of ~E13.5 testis (c, bar = 100 μm; bar = 50 μm; e, bar = 20 μm). Three out of 12 testes analysed had I-1-positive sex cords at this age, but older gonads showed no internal expression of th protein, suggesting down-regulation of I-1 after mesothelial cell migration.

.cknowledgements

This work was funded by the Sir Stanley and Lady David-on Trust, and the University of Edinburgh Moray Endow-ıent Fund. We would like to thank Linda Sharp for her help /ith the confocal microscopy, and Dr Jamie Davies and .ndrew J. Wilson for helpful discussion and comments on ıe manuscript.

.eferences

itken, A., Bilham, T., Cohen, P., 1982. Complete primary structure of protein phosphatase inhibitor-1 from rabbit skeletal muscle. Eur. J. Biochem. 126, 235–246.

ressler, G.R., Douglass, E.C., 1992. Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumour. Proc. Natl. Acad. Sci. USA 89, 1179–1183.

lbrecht, A., DiRenzo, J., Smith, R.G., Shenolikar, S., 1990. Molecular cloning of protein phosphatase inhibitor-1 and its expression in rat and rabbit tissues. J. Biol. Chem. 265, 13415–13418.

Endo, S., Zhou, X., Connor, J., Wang, B., Shenolikar, S., 1996. Multiple structural elements define the specificity of recombinant human inhibitor-1 as a protein phosphatase-1 inhibitor. Biochemistry 35, 5220–5228.

Fantes, J.A., Bickmore, W.A., Fletcher, J.M., Ballesta, F., Hanson, I.M., Van Heyningen, V., 1992. Submicroscopic deletions at the WAGR locus, revealed by nonradioactive in situ hybridization. Am. J. Hum. Genet. 51, 1286–1294.

Gardiner-Garden, M., Frommer, M., 1987. CpG islands in vertebrate genomes. J. Mol. Biol. 196, 261–282.

Lennon, G., Auffray, C., Polymeropoulos, M., Soares, M.B., 1996. The IMAGE consortium: an integrated molecular analysis of genomes and their expression. Genomics 33, 151–152.

Svennilson, J., Durbeej, M., Celsi, G., Laestadius, A., da Cruz e Silva, E.F., Ekblom, P., Aperia, A., 1995. Evidence for a role of protein phosphatases 1 and 2A during early nephrogenesis. Kidney Int. 48, 103–110.

Williams, P.L., Bannister, L.H., Berry, M.M., Collins, P., Dysan, M., Dussek, J.E., Ferguson, M.W.J., (Eds.). Gray's Anatomy, 38th edn., Churchill Livingstone, Edinburgh, 1995, pp. 67–75.